

AD-A101 802

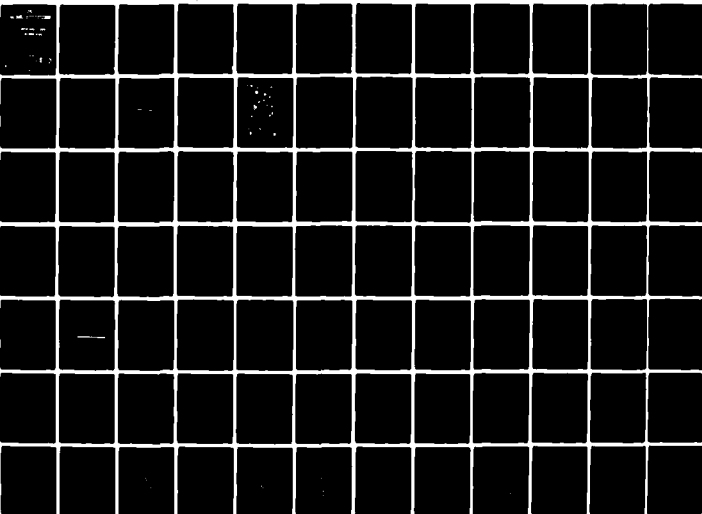
CORPS OF ENGINEERS WASHINGTON DC
PROCEEDINGS OF A SEMINAR ON WATER QUALITY DATA INTERPRETATION, --ETC(U)
1978

F/G 13/2

UNCLASSIFIED

NL

1 of 3
AD-A101 802



LEVEL II

(13)



COMMITTEE ON WATER QUALITY

AD A101802

PROCEEDINGS OF A SEMINAR ON
**WATER QUALITY DATA
INTERPRETATION**

WATER QUALITY DATA

8-9 FEB 1978
ATLANTA, GEORGIA

DTIC
ELECTE
JUL 23 1981
S H D

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

81 7 22 0 79

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|--|-----------------------|--|
| 1. REPORT NUMBER | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| | AD-A402 802 | |
| 4. TITLE (and Subtitle) | | 5. TYPE OF REPORT & PERIOD COVERED |
| PROCEEDINGS OF SEMINAR ON WATER QUALITY DATA INTERPRETATION | | |
| 6. PERFORMING ORG. REPORT NUMBER | | |
| 7. AUTHOR(s) | | 8. CONTRACT OR GRANT NUMBER(s) |
| Committee on Water Quality Collection of 10 individual papers edited by R. G. Willey. | | |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| US Army Corps of Engineers The Hydrologic Engineering Center 609 Second Street, Davis, CA 95616 | | |
| 11. CONTROLLING OFFICE NAME AND ADDRESS | | 12. REPORT DATE |
| | | 8-9 February 1978 |
| | | 13. NUMBER OF PAGES |
| | | 193 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 15. SECURITY CLASS. (of this report) |
| | | Unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |
| 16. DISTRIBUTION STATEMENT (of this Report) | | |
| Distribution of this publication is unlimited. | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES | | |
| Seminar for Corps of Engineers Staff held in Atlanta, Georgia on 8-9 February 1978. | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) | | |
| Water Quality, Data Interpretation, Aquatic Biology, Statistics, Data Storage System, Data Collection. | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) | | |
| Topics addressed during the seminar include data collection, data processing, laboratory quality control and interpretation of physical-chemical, biological, and chemical contaminant data. | | |

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

...at the ... laboratory and interpretation of physical-chemical ...
... data collection, data processing, and statistical analysis ...
... laboratory and interpretation of physical-chemical ...

... data collection, data processing, and statistical analysis ...
... laboratory and interpretation of physical-chemical ...

... data collection, data processing, and statistical analysis ...
... laboratory and interpretation of physical-chemical ...

... data collection, data processing, and statistical analysis ...
... laboratory and interpretation of physical-chemical ...
... data collection, data processing, and statistical analysis ...
... laboratory and interpretation of physical-chemical ...

INSTRUCTIONS FOR PREPARATION OF REPORT DOCUMENTATION PAGE

RESPONSIBILITY. The controlling DoD office will be responsible for completion of the Report Documentation Page, DD Form 1473, in all technical reports prepared by or for DoD organizations.

CLASSIFICATION. Since this Report Documentation Page, DD Form 1473, is used in preparing announcements, bibliographies, and data banks, it should be unclassified if possible. If a classification is required, identify the classified items on the page by the appropriate symbol.

COMPLETION GUIDE

General. Make Blocks 1, 4, 5, 6, 7, 11, 13, 15, and 16 agree with the corresponding information on the report cover. Leave Blocks 2 and 3 blank.

Block 1. Report Number. Enter the unique alphanumeric report number shown on the cover.

Block 2. Government Accession No. Leave blank. This space is for use by the Defense Documentation Center.

Block 3. Recipient's Catalog Number. Leave blank. This space is for the use of the report recipient to assist in future retrieval of the document.

Block 4. Title and Subtitle. Enter the title in all capital letters exactly as it appears on the publication. Titles should be unclassified whenever possible. Write out the English equivalent for Greek letters and mathematical symbols in the title (see "Abstracting Scientific and Technical Reports of Defense-sponsored RDT/E," AD-667 000). If the report has a subtitle, this subtitle should follow the main title, be separated by a comma or semicolon if appropriate, and be initially capitalized. If a publication has a title in a foreign language, translate the title into English and follow the English translation with the title in the original language. Make every effort to simplify the title before publication.

Block 5. Type of Report and Period Covered. Indicate here whether report is interim, final, etc., and, if applicable, inclusive dates of period covered, such as the life of a contract covered in a final contractor report.

Block 6. Performing Organization Report Number. Only numbers other than the official report number shown in Block 1, such as series numbers for in-house reports or a contractor/grantee number assigned by him, will be placed in this space. If no such numbers are used, leave this space blank.

Block 7. Author(s). Include corresponding information from the report cover. Give the name(s) of the author(s) in conventional order (for example, John R. Doe or, if author prefers, J. Robert Doe). In addition, list the affiliation of an author if it differs from that of the performing organization.

Block 8. Contract or Grant Number(s). For a contractor or grantee report, enter the complete contract or grant number(s) under which the work reported was accomplished. Leave blank in in-house reports.

Block 9. Performing Organization Name and Address. For in-house reports enter the name and address, including office symbol, of the performing activity. For contractor or grantee reports enter the name and address of the contractor or grantee who prepared the report and identify the appropriate corporate division, school, laboratory, etc., of the author. List city, state, and ZIP Code.

Block 10. Program Element, Project, Task Area, and Work Unit Numbers. Enter here the number code from the applicable Department of Defense form, such as the DD Form 1498, "Research and Technology Work Unit Summary" or the DD Form 1634, "Research and Development Planning Summary," which identifies the program element, project, task area, and work unit or equivalent under which the work was authorized.

Block 11. Controlling Office Name and Address. Enter the full, official name and address, including office symbol, of the controlling office. (Equates to funding/sponsoring agency. For definition see DoD Directive 5200.20, "Distribution Statements on Technical Documents.")

Block 12. Report Date. Enter here the day, month, and year or month and year as shown on the cover.

Block 13. Number of Pages. Enter the total number of pages.

Block 14. Monitoring Agency Name and Address (if different from Controlling Office). For use when the controlling or funding office does not directly administer a project, contract, or grant, but delegates the administrative responsibility to another organization.

Blocks 15 & 15a. Security Classification of the Report: Declassification/Downgrading Schedule of the Report. Enter in 15 the highest classification of the report. If appropriate, enter in 15a the declassification/downgrading schedule of the report, using the abbreviations for declassification/downgrading schedules listed in paragraph 4-207 of DoD 5200.1-R.

Block 16. Distribution Statement of the Report. Insert here the applicable distribution statement of the report from DoD Directive 5200.20, "Distribution Statements on Technical Documents."

Block 17. Distribution Statement (of the abstract entered in Block 20, if different from the distribution statement of the report). Insert here the applicable distribution statement of the abstract from DoD Directive 5200.20, "Distribution Statements on Technical Documents."

Block 18. Supplementary Notes. Enter information not included elsewhere but useful, such as: Prepared in cooperation with . . . Translation of (or by) . . . Presented at conference of . . . To be published in . . .

Block 19. Key Words. Select terms or short phrases that identify the principal subjects covered in the report, and are sufficiently specific and precise to be used as index entries for cataloging, conforming to standard terminology. The DoD "Thesaurus of Engineering and Scientific Terms" (TEST), AD-672 000, can be helpful.

Block 20. Abstract. The abstract should be a brief (not to exceed 200 words) factual summary of the most significant information contained in the report. If possible, the abstract of a classified report should be unclassified and the abstract to an unclassified report should consist of publicly-releasable information. If the report contains a significant bibliography or literature survey, mention it here. For information on preparing abstracts see "Abstracting Scientific and Technical Reports of Defense-Sponsored RDT&E," AD-667 000.

6

PROCEEDINGS OF A SEMINAR

ON

WATER QUALITY DATA INTERPRETATION

8-9 February 1978
Atlanta, Georgia

12) 298/

11 2978

U.S. ARMY CORPS OF ENGINEERS

COMMITTEE ON WATER QUALITY

WASHINGTON, D.C.

| | |
|--------------------|-------------------------------------|
| Accession For | |
| NTIS GRA&I | <input checked="" type="checkbox"/> |
| DTIC TAB | <input type="checkbox"/> |
| Unannounced | <input type="checkbox"/> |
| Justification | |
| By | |
| Distribution/ | |
| Availability Codes | |
| Dist | Avail and/or Special |
| A | |

099300

FOREWORD

A two-day seminar on Water Quality Data Interpretation was held in Atlanta, Georgia on 8-9 February 1978. ^{the} The purpose of ~~the~~ seminar was to provide a forum for Corps of Engineers personnel who are routinely involved in water quality data collection, analysis, and interpretation. Topics addressed ~~during the seminar~~ include data collection, data processing, laboratory quality control, and interpretation of physical-chemical, biological, and chemical contaminant data. Ten papers presented during the seminar are ~~contained herein~~. *provided.*

The views and conclusions expressed in these proceedings are those of the authors and are not intended to modify or replace official guidance or directives such as engineer regulations, manuals, circulars, or technical letters issued by the Office of the Chief of Engineers.

R. G. WILLEY
Editor

CONTENTS

Paper Number

THE COMMITTEE ON WATER QUALITY

Janice E. Rasgus
Vice Chairman, Committee on Water Quality
North Central Division 1

ENVIRONMENTAL AND WATER QUALITY OPERATIONAL STUDIES

Milton Millard
Chief Western Section Operations Branch
Con-Ops Division
Directorate of Civil Works
Office Chief of Engineers 2

WATER QUALITY DATA COLLECTION

Earl E. Eiker
Hydrologic Engineering Section
Office Chief of Engineers 3

OHIO RIVER DIVISION'S LABORATORY MASTER FILE SYSTEM

Henry C. Jackson
Ohio River Division
Reservoir Control Center
Water Quality Section 4

STATISTICAL INTERPRETATION OF WATER QUALITY DATA

James L. Grant
Nuclear Engineering Company, Inc., Louisville, KY
Formerly Chief Hydrologist
Law Engineering Testing Company 5

DATA VALIDATION

James D. Ashworth
Water Quality Section
Huntington District 6

TEMPERATURE AND DISSOLVED OXYGEN DATA INTERPRETATION

Dennis E. Ford
Environmental Laboratory
U.S. Army Engineer Waterways Experiment Station
Hydrologist 7

Paper Number

**COLIFORMS AND WATER QUALITY: USE OF
DATA IN PROJECT DESIGN AND OPERATION**

Kent W. Thornton
Environmental Laboratory
U. S. Army Engineer Waterways Experiment Station
Research Ecologist 8

**AQUATIC BIOLOGY - DATA INTERPRETATION (PART I):
PRIMARY PRODUCER ORGANISMS (ALGAE AND MACROPHYTES)**

John W. Barko
Environmental Effects Laboratory
Waterways Experiment Station
Research Biologist 9

**WATER QUALITY EFFECTS OF MULTI-PURPOSE RESERVOIRS
AND NAVIGATION DAMS ON THE MONONGAHELA RIVER**

Michael Koryak
Hydrology and Hydraulics Branch
U. S. Army Engineer District Pittsburgh
Aquatic Biologist 10

SEMINAR ON
WATER QUALITY DATA INTERPRETATION

8-9 February 1978
Atlanta, Georgia

Committee Members

| | |
|---------------------|--|
| Dr. Mark Anthony | Ohio River Division |
| Charles Bradshaw | Lower Mississippi Valley Division |
| John Bushman | Office of the Chief of Engineers Planning Division |
| Dick DiBuono | New England Division |
| Harry Dotson | South Pacific Division |
| Earl Eiker | Office of the Chief of Engineers Engineering Division |
| Dr. Rex Eley | Waterways Experiment Station Environmental Effects Laboratory |
| John Grace | Waterways Experiment Station Hydraulics Laboratory |
| Al Harrison | Missouri River Division |
| Dave Legg | North Pacific Division |
| Tom Maisano | North Atlantic Division |
| Dr. Harlan L. McKim | Cold Regions Research and Engineering Laboratory |
| Milt Millard | Office of the Chief of Engineers Construction-Operations Division |
| Ms. Jan Rasgus | North Central Division |
| Julian Raynes | South Atlantic Division |
| Charles Sullivan | Southwestern Division |
| R. G. Willey | The Hydrologic Engineering Center |

Seminar Speakers

| | |
|---------------------|---|
| Mr. Jim Ashworth | Huntington District |
| Dr. John Barko | Waterways Experiment Station Environmental Effects Laboratory |
| Mr. Earl Eiker | Office of the Chief of Engineers Engineering Division |
| Dr. Bob Engler | Waterways Experiment Station Environmental Effects Laboratory |
| Dr. Dennis Ford | Waterways Experiment Station Environmental Effects Laboratory |
| Mr. Jim Grant | Law Engineering Testing Company |
| LTC John Hill | Office of the Chief of Engineers Civil Works, Environmental Programs |
| Mr. Henry Jackson | Ohio River Division |
| Dr. Bob Kennedy | Waterways Experiment Station Environmental Effects Laboratory |
| Mr. Mike Koryak | Pittsburgh District |
| BG Kenneth McIntyre | South Atlantic Division |
| Mr. Milt Millard | Office of the Chief of Engineers Construction Operations Division |
| Ms. Jan Rasgus | North Central Division |
| Dr. Kent Thornton | Waterways Experiment Station Environmental Effects Laboratory |
| Mr. Jim Westhoff | Waterways Experiment Station Environmental Effects Laboratory |

THE COMMITTEE ON WATER QUALITY

By

Janice E. Rasgus¹

The Committee on Water Quality was formally established by ER 15-2-10, dated 15 December 1975. The members of the Committee are appointed by the Chief, Engineering Division, Directorate of Civil Works, Office, Chief of Engineers and represent each Division and several of the laboratories. The Committee is charged with providing guidance for developing a comprehensive, coordinated Corps-wide water quality management program.

The Committee met three times during 1977 and accomplished the following:

- 1) Recommended three water quality training courses which will be held during FY 78. These are:
 - a. Selective Withdrawal Design.
 - b. Water Quality and Ecosystem Modeling.
 - c. Water Quality Aspects of Water Control.
- 2) Provided a preliminary listing of water quality/environmental training courses available to Corps personnel through other agencies.
- 3) Developed a workshop for laboratory chemists at WES to be held on 22 and 23 February 1978. The purpose of the workshop is to discuss the technical aspects of laboratory quality control and analytical procedures.
- 4) Sponsored a water quality seminar, "Water Quality Data Collection and Management", which was held in Denver, Colorado, on 25-26 January 1977. A proceedings was published and distributed to each participant.
- 5) Revised and issued ER 1130-2-334, dated 16 December 1977. It requires the reporting of water quality activities.
- 6) Drafted an ER which addresses quality assurance procedures for Corps and contracted laboratories and laboratory personnel needs.
- 7) Issued ER 15-2-11, dated 1 November 1977. It describes the Committee's consulting service and the procedures for obtaining these services. This service is one of the Committee's important functions and is desirable for disseminating information and providing a means for effective Corps-wide communication.

¹ North Central Division; Vice Chairman, Committee on Water Quality

8) Consulted with the Louisville District on the Cave Run reservoir project in Morehead, Kentucky. During the summer months the reservoir experiences low dissolved oxygen concentrations and consequent releases of slugs of high concentrations of iron, manganese and hydrogen sulfide. The meeting helped generate a better understanding of the problems and developed some practical solutions.

During the coming year the Committee will address the following items:

- 1) Manpower requirements for properly staffing a water quality management function.
- 2) Data management systems.
- 3) Water quality laboratory quality control and personnel needs.
- 4) Technology transfer.
- 5) Training needs.

Members of the Committee should be contacted with any questions, concerns or suggestions. These will be brought to the attention of the Committee for discussion. The members of the Committee on Water Quality are:

Dr. Mark Anthony - ORD
Mr. Charles Bradshaw - LMVD
Mr. John Bushman - OCE
Mr. Richard DiBuono - NED
Mr. Harry Dotson - SPD
Mr. Earl Eiker - OCE
Dr. Rex Eley - WES
Mr. John Grace - WES
Mr. Alfred Harrison - MRD
Mr. David Legg - NPD
Mr. Thomas Maisano - NAD
Dr. Harlan McKim - CRREL
Mr. Milton Millard - OCE
Ms. Janice Rasgus - NCD
Mr. Julian Raynes - SAD
Mr. Charles Sullivan - SWD
Mr. Jerry Willey - HEC

ENVIRONMENTAL AND WATER QUALITY

OPERATIONAL STUDIES

by

Milton Millard *

The Civil Works Program of the Corps is unique as it encompasses the entire spectrum of Water Resource Development. Traditionally, the role has involved the planning, design, construction, and operations of water resources, primarily for navigation, flood control, hydropower and associated activities. The recent concern for the preservation and protection of the environment has resulted in legislation that makes environment and water quality additional considerations in the management of these water resources and has added new dimensions to our basic management requirements. For projects in the planning phase, this means that the environmental and water quality objectives must be addressed along with the basic project purposes. For projects that are operational, environmental and water quality objectives mean additional pressures to meet new standards and reduce adverse environmental impacts. To help meet these new responsibilities, the Corps has undertaken another major research effort.

Without going into detail on the background of the study, the sequence of events depicted (Fig. 1) followed an OCE staff study made in 1975. Problem identification and assessment was accomplished in 1976, detailed program planning was done in 1977 and the research and study program is being initiated in FY 1978. For budget purposes, the program was identified as "The Environmental-Water Quality Operational Studies Program: and was immediately referred to by the acronym, "EWQOS."

Environmental quality problem identification and assessment was the first step in formulation of the EWQOS program. A field office survey in which all divisions were involved was conducted by a team from OCE and WES. The major purpose of this survey was to obtain information on the impacts of the new requirements on the various functional areas of the Corps' organization, namely, the Planning, Engineering, and Construction-Operations Divisions. Some of the impacts of Environmental Quality problems on Civil Works activities are shown in Figure 2.

Environmental and Water Quality Problems were documented at 152 Corps Projects representing 200 megawatts of power generating capacity, 300 billion gallons of water supply, and over 125 million recreation days annually. A major finding was that the environmental and water quality problems most significantly impact on Project Operations and are not limited to any one part of the country.

* Milton Millard, Chief Western Section, Operations Branch, Con-Ops Division, Directorate of Civil Works, Office Chief of Engineers

Figure 3 gives some idea of how the Corps' projects are distributed around the country, but does not necessarily indicate the distribution of the major environmental problems.

The end product of the first phase of this program was a report that served as justification for the program to OMB and the Congress. A few of the problems identified and questions raised by the report are:

- a. What effect do Corps projects have on water quality, is the effect significant, and is there a need to find a solution?
- b. Under the present energy situation and the increased importance placed on hydropower, what are the effects of hydropower operations at pumpback facilities?
- c. What do we know of the quality of reservoir releases and their impact downstream?
- d. What are the impacts of flow variations from reservoir projects? What are the minimum flow requirements? What are the impacts of hydropower peaking operations and other flow variations?
- e. What are the impacts of fluctuating reservoir pool elevations, i.e., exposed mud flats, aesthetics, the effects on fisheries and other habitats?
- f. What about reservoir eutrophication and problems of algal blooms? How does this affect project purposes such as water supply or recreation?

The report also indicates a need to improve procedures in the preparation of environmental assessments, a need for definition of data requirements and improved data management techniques. In addition, the report points out the need for a better understanding of the impacts of the various activities in and on the waterways on the environment. What, if any, adverse impacts are there from navigation, from the construction of cut-offs and training structures, bank erosion control methods, and the numerous other activities?

The formulation of the EWQOS Program considered the problems, the possible resources, the time to be available for the work, and the following concepts:

- A. The probability of Success
- B. Technical constraints
- C. Cost of the Research
- D. Potential solutions

The program, as developed, is organized into six research projects plus associated field studies aimed at meeting the program objectives. These field studies form a major part of the program. Examples of major research areas to be addressed are:

- a. The improved understanding of reservoir hydrodynamics through the use of scaled models to get a better feeling for reservoir project effects on water quality and ecology.
- b. Control of nuisance algal blooms - what are the major factors causing them and what steps may be taken to control them?
- c. Reservoir Release Criteria - what is required to maintain a suitable downstream habitat how may the effects of algal releases be reduced?

The success of this program is dependent on the interactions of all elements within the Corps and the many outside interests involved in environmental and water quality problems.

In the area of technology transfer and application, we will try to be innovative and expand from the usual scientific reports that usually do not get to the field operating people. In each report, when practicable, we will include a "How to Apply" section or appendix which will also be published as a technical letter or manual. In cases where extensive research work units are involved, we will develop synthesized reports and "How To" manuals. In addition, where practical, we will provide field demonstrations of new techniques and procedures.

To assist in this, each District and Division will be expected to provide continuous feedback on its research need and the results of the program's outputs. To facilitate this feedback and assure that your needs are fed into the program, a Field Review Group has been established to work with the technical monitors in management of the program. You should all become acquainted with your division representative on the Field Review Group and work closely with him to assure that your needs are considered.

To get the most out of this program and avoid duplication of efforts, all activities have been and will continue to be coordinated closely with Federal and State agencies, consultants from industry and academia, and with private interest groups. Early in the program we held an interagency briefing on our proposed activities and received valuable feedback. We propose another such briefing in the near future and then on a regular schedule as the program progresses.

The EWQOS Program is based on a six-year effort at an anticipated cost of \$30 million. This is, of course, a planning figure which will be reviewed annually by the Administration and the Congress.

The initial planning for the program has been completed - we are funded for our first year's efforts. Field sites have been selected and work will be initiated after the selections are formally approved by the Division Engineers involved. In a short time, the first RFP will

go out and research efforts will be moving along at WES. To put this all together and assure success, we, the Technical Monitors and OCE, solicit your support.

Figure 1.
SEQUENCE OF EVENTS
LEADING TO EWQOS

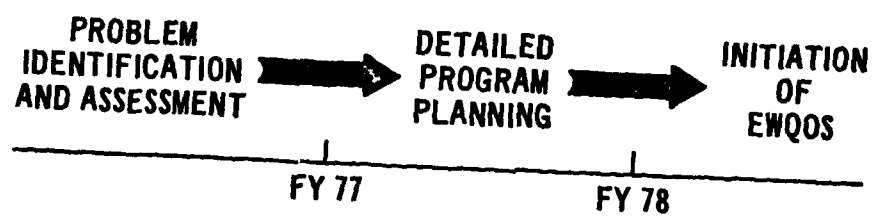


Figure 2.

**IMPACTS OF ENVIRONMENTAL QUALITY PROBLEMS
ON CIVIL WORKS ACTIVITIES OF THE CORPS**

| FUNCTION | IMPACT |
|---------------------|--|
| PLANNING | INCREASED COSTS FOR ASSESSMENT LITIGATIONS ASSOCIATED WITH ENVIRONMENTAL ASSESSMENT |
| ENGINEERING | DESIGN MODIFICATIONS |
| CONSTRUCTION | COST ESCALATION DUE TO DELAYS ADDITIONAL CONSTRUCTION TO SOLVE PROBLEM |
| OPERATION | LOSS OF PROJECT BENEFITS INCREASED OPERATIONAL COSTS |

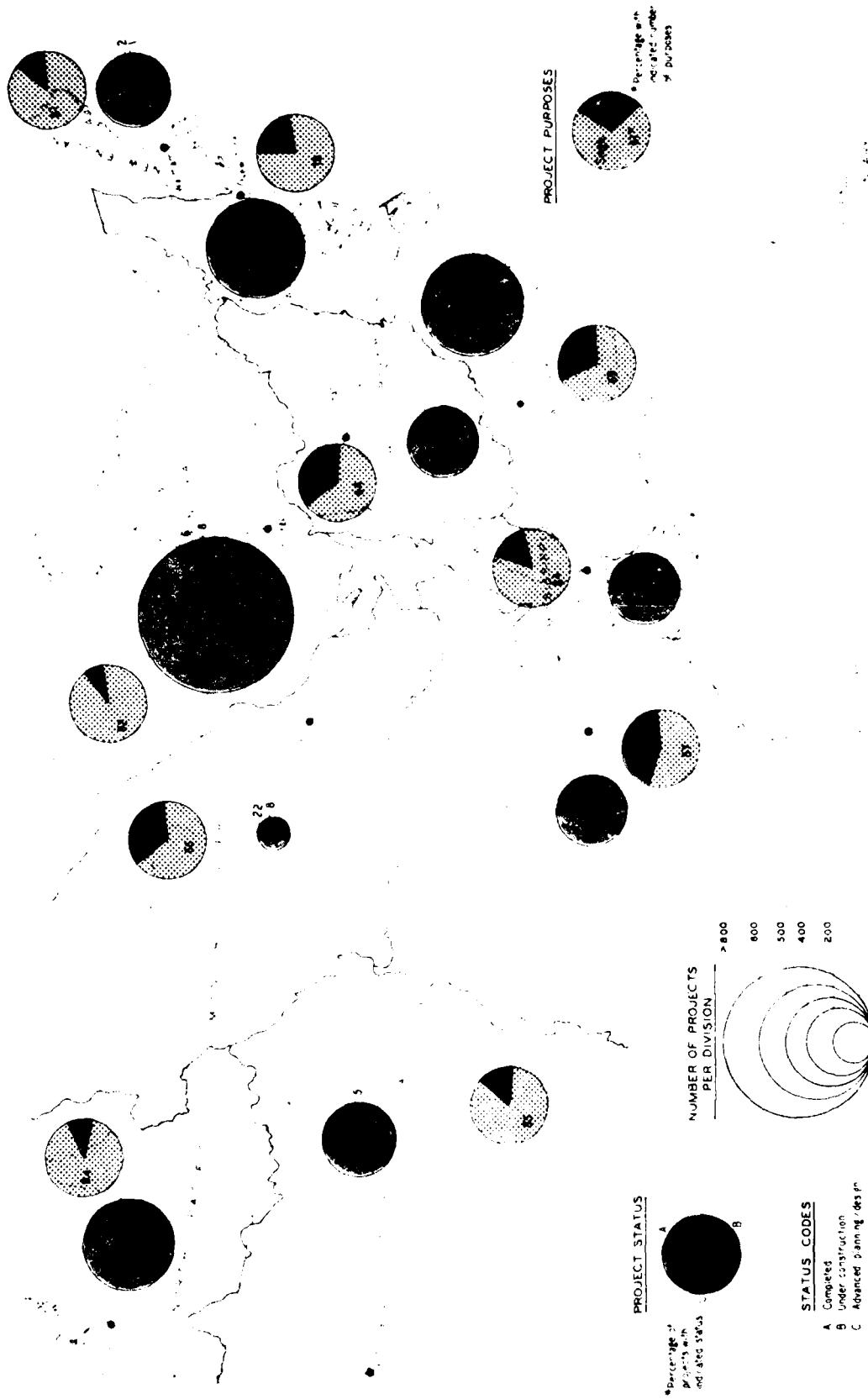


Figure 3. Status and number of purposes of Civil Works projects

WATER QUALITY DATA COLLECTION

BY

1

EARL E. EIKER

Over the last ten years there has been a dramatic increase in the interest in water quality problems and in finding solutions to these problems. At the same time traditional Corps roles in water resources development and management have magnified these concerns relative to our activities. Water quality studies are now conducted at every phase of project development; survey reports, general and feature design memoranda, and for projects that are completed and those under construction. Each water quality study is different; yet at the same time there is a common denominator, and the common denominator is data. You simply can't do a water quality study without data. At the Denver Seminar last year, two days were spent talking about data collection, and little more was done than scratching the surface.

Someone might ask the question, why should the Corps collect water quality data? The states are collecting data, EPA is collecting data, USGS, USFWS, and universities are all collecting data for one purpose or another. But that's just the point, their purposes for collecting data are considerably different in the vast majority of cases from ours. For example, when EPA collects data to monitor industrial discharges, 99% of the time that data will not be of any worth to a person sitting in a District office trying to regulate a reservoir. This leaves us in a position where for Corps unique purposes, at best we must collect data to supplement that available from others and at worst we must conduct complete data collection programs of our own.

While looking back over approximately 10 years of notes related to data collection, I came across an interesting piece from a 1969 training course that was held at HEC. I think it points out some problems that still exist with data collection. It is not always possible to anticipate future needs, hence there exists a tendency for the creation of data programs for data sake alone, ignoring the purposes to which the information gathered should be directed. Regretably, there too often seems to be money to gather more data, but relatively little to evaluate it. Thus, the process goes on at ever expanding levels until we become virtually inundated with data. At this juncture, we may have to give some attention to data disposal. At the very least we should look toward optimizing our efforts in data management so that we get maximum value for each dollar expended. While we have made improvements, too often we wind up in the same situation today as we were in 10 years ago. We collect too much data and often not the right data. We need to get only the information we need and then move on to interpret it and come up with the solutions to the problems.

1

Hydrologic Engineering Section, Office Chief of Engineers

Paper 3

The challenge of developing and managing data collection programs is great. It taxes the ingenuity of every water quality manager. A major difficulty in developing efficient data collection efforts is that "cookbook" approaches to sampling program design simply can't be set up. A quick look at why we collect water quality data will illustrate this point. Data are presently collected to provide project baseline information, for early identification of problems, to provide guidance to reservoir regulation elements, as a basis for development of design criteria, as input to EIS's and for numerous other reasons. The needs are clearly different and the programs to meet these needs must be just as diverse. Under all this, however, there is a logical general approach to sampling program design and management which, if followed, will produce good results.

For a water quality data collection program to be successful, many important factors must be brought together. The program must first be established to meet an identified need. It must address only very well defined objectives which have been developed based on the need. It has to be a team effort in the sense that it must include input from both field and laboratory personnel in order to reflect reasonable workloads and minimize the possibility of technical problems arising during the implementation phase. Finally, the output from the effort must satisfy the need that the program was designed to address.

Figure 1 illustrates the interactions required to set up a data collection program. The data collection program is in the center. We have needs considered, a program manager, and input from field and lab personnel. Also, output, feedback and constraints are shown.

Now let's take these items one at a time. There are probably an infinite number of needs for which to collect water quality data. Some are to develop baseline information, to assess our projects, as a basis for project modification, for early problem identification, for operations and to provide reservoir regulation guidance. We collect data for criteria development, for design, for EIS preparation. We collect data for use in coordination and cooperation with other agencies, Federal and state. We use their data whenever we can and they use our data whenever they can. We need data in order to administer the regulatory program. We need to set up the program to match the need.

Next, there is the program manager. What's his job? He first must assimilate the needs and establish the objectives. He is responsible for setting up and conducting the program. To do this he must assess laboratory and field capability (with suitable input from these people) in order to determine whether he uses inhouse people or contracts the effort. After the program is underway he receives the data and is responsible for data management. He adapts the program to meet changing objectives and needs. He's responsible for interpreting the data. He is responsible for the output which is a continuous process. He has to encourage feedback. This is the only way that we are going to be able to keep up the interest of the laboratory and field people. He must let them know what the data is used for and why it is needed.

What are the constraints? The most critical one is manpower. There never seems to be enough people to do all the work. Another is time. Invariably the need for the data and the need for the evaluation of the data was "yesterday" so we are faced with a situation where we must set up a program that is greatly compressed in time. The final constraint is money. But keep in mind we need a good balance between data collection and evaluation. Both are needed to make a study successful. Data collected but not evaluated should not have been collected in the first place.

What about lab personnel? One of the most important considerations here is the need for laboratory people to input certain technical considerations to the program manager so he can be fully aware of the problems the laboratory may encounter over the life of the program. Input such as required level of detection for given parameters and the capability of the lab as it relates to volume, workload and what analyses the lab is capable of performing. There is no point in setting up a lab program if your laboratory people don't have the capability of doing it. Finally the lab personnel can give you some insight as to "turnaround time" you can reasonably expect.

What about field personnel? The same considerations apply to field personnel. If you've got a field crew to collect the data, their technical input and knowledge of the area should be a guide in setting up a program; particularly when it comes to frequency of sampling, location of sampling stations, etc. You can't get away from the fact that the man who knows the area can provide you with a lot of information. This aspect is too often neglected in setting up a program.

Finally let's consider the output. This item not only represents the final report but also the continual interpretation and feedback of data. We must continually evaluate our data so that we can make improvements and adjustments to the program. Output also includes data management whether it's a file drawer, a computer system or whatever.

Now how do we pull all this together? Figure 2 shows five basic steps needed to establish an effective data collection program. The first step, and by far the most critical, is to define the scope and objectives to meet the identified needs. After the objectives are set we must assemble all the available data. This includes any pertinent data collected by others which will often give some indication of what we need to collect and how to collect it. Information such as available water quality data, land use information, point and non-point sources of pollution in the watershed and maps of the area is pertinent. All this information should be assembled and evaluated before the plan is put together. Thirdly, a preliminary program is established. This is the point where the laboratory and the field personnel should be brought into the process to make it a coordinated effort. A good idea is to keep the preliminary plan unconstrained for the first cut. Since we never have enough money, manpower and time to do all we would like to do, developing an unconstrained program first will give us the base to evaluate the effect of reductions when the constraints are applied.

In other words, what we are losing as we reduce the scope of the program to meet funding, personnel and time limitations. The fourth item is to review the preliminary plan, possibly collecting some water quality data to support out initial judgements. This review enables us to then develop the final program. Ask yourself again. Does the program address the objectives? Is the program going to meet the needs? Is it going to produce the output required? Is the program cost effective? If the answer to these questions is yes, we now have our final program.

Now for the data collection program itself. Figure 3 shows seven items common to any water quality data collection program. The first one is the objectives. Make them specific - define them well and you'll save yourself a lot of problems as you get into the program. The second thing is parameters. What do we really need to know? It may be all right in the first cut to take a large number of samples and analyze many parameters. However, you better come back to the office, do some interpretation and think about cutting down. Don't be drawn into the situation where you are collecting a lot of information because someone five years from now may have a use for it. The third item is sampling sites. These have to be representative. There's no use sampling unless it gives you a good picture of what the water quality is. The fourth item is sampling frequency. Get enough samples to answer the questions but not too many. This item has the biggest potential for wasted effort. Often data is collected more frequently than we really need to answer the questions. Next is sample handling and preservation. What are the requirements? This is where the experience of the laboratory and field people can be of great assistance. The sixth item is the lab and field analyses. What's needed? Should the effort be contracted or done in house. This should be clearly laid out. Along these same lines, resist the temptation to put water quality monitors here, there and everywhere. They're often expensive to run, have operation and maintenance problems and yield more data than we need. The final item is flexibility. We should be constantly reviewing the program and making changes. If the program can't be modified with relative ease our task will be much more difficult.

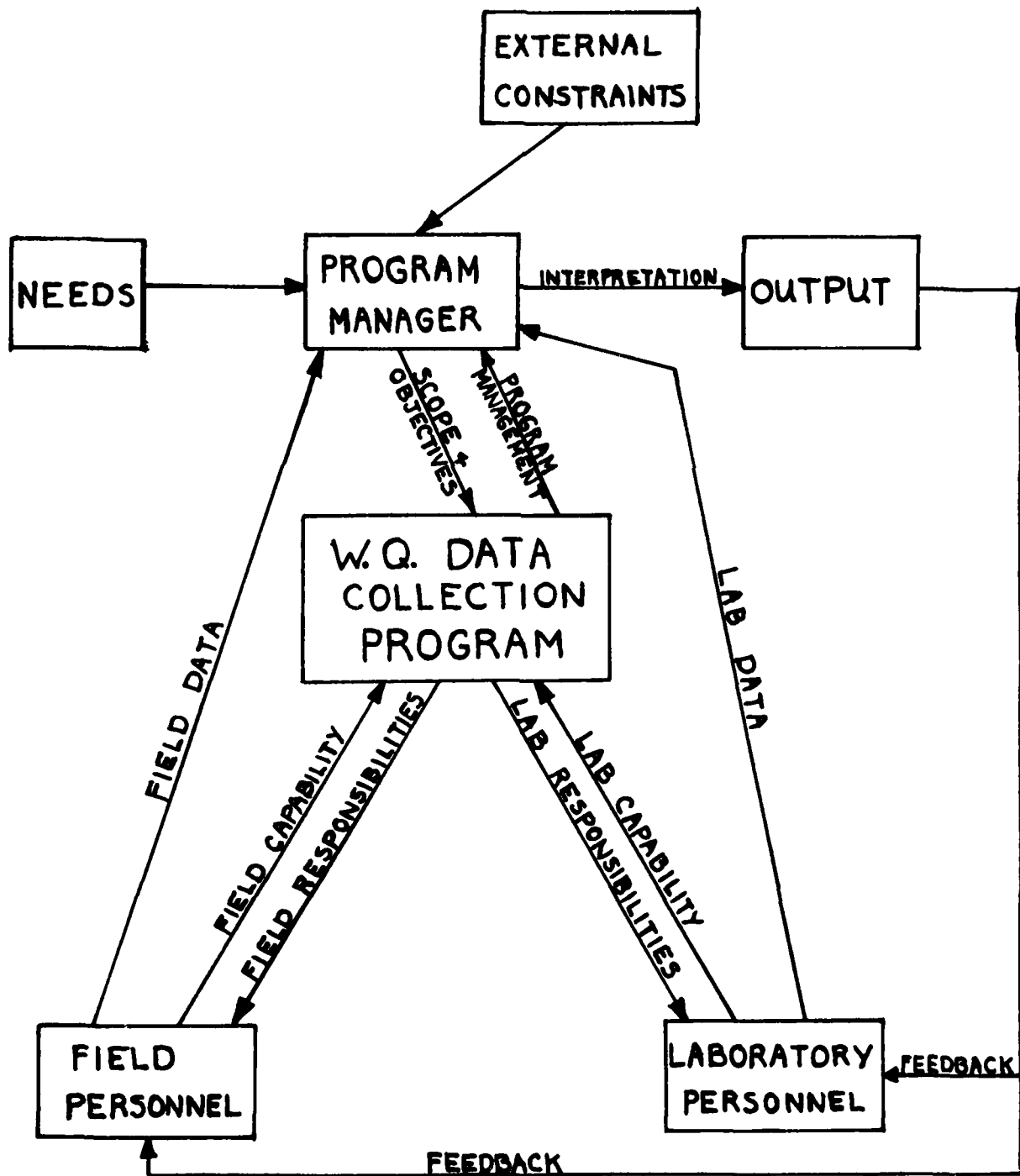
Factors to be considered under each of the seven items discussed above are shown on Figures 4-10. While the figures may not be all inclusive, they should nevertheless provide a guide to setting up most water quality sampling programs.

In summary, developing and managing water quality data collection programs is not an easy task. It requires co-ordination and communication with many interested parties. Furthermore, there is no standard procedure to follow. However, if the ideas presented in this paper are followed, we should be able to measurably improve our efforts in this important area.

References

1. Committee on Water Quality, "Water Quality Data Collection and Management," Seminar Proceedings, 25-26 January 1977, Denver, Colorado, U.S. Army Corps of Engineers, Washington, D.C., 1977.
2. Training Course, Number 27, Water Quality Management, HEC, Davis, California, March 1969.

DATA COLLECTION PROGRAM FLOW CHART



DATA COLLECTION
PROGRAM DEVELOPMENT

1. ASSIMILATE NEEDS & ESTABLISH SCOPE AND OBJECTIVES
2. ASSEMBLE ALL AVAILABLE INFORMATION
3. ESTABLISH PRELIMINARY PLAN
4. REVIEW PLAN
5. ESTABLISH FINAL PLAN

DATA COLLECTION
PROGRAM

1. OBJECTIVES
2. PARAMETERS
3. SAMPLING SITES
4. SAMPLING FREQUENCY
5. SAMPLE HANDLING & PRESERVATION
6. LAB & FIELD ANALYSES
7. FLEXIBILITY

FACTORS TO CONSIDER
IN SETTING OBJECTIVES

1. NEED FOR DATA
2. PHASE OF PROJECT
3. TYPE OF PROJECT
4. CONSTRAINTS
5. PROBABILITY OF SUCCESS

FACTORS TO CONSIDER
IN PARAMETER SELECTION

1. RELATIONSHIP TO OBJECTIVES
2. WATERSHED LAND USES
3. PROJECT PURPOSES
4. KNOWN OR ANTICIPATED PROBLEMS
5. COST OF ANALYSES
6. SAMPLE PRESERVATION REQUIREMENTS
7. LABORATORY SUPPORT CAPABILITY
8. SECONDARY DATA REQUIREMENTS

FACTORS TO CONSIDER IN
SAMPLING SITE SELECTION

1. RELATIONSHIP TO OBJECTIVES
2. FLOW, CURRENT AND MIXING PATTERNS
3. PHYSICAL CHARACTERISTICS OF WATERBODY
4. POLLUTANT SOURCES
5. ACCESSABILITY OF SITE
6. ECONOMICS

FACTORS TO CONSIDER
IN SAMPLING FREQUENCY

1. RELATIONSHIP TO OBJECTIVES
2. NORMAL VARIATIONS IN CONCENTRATION OF PARAMETERS
3. HYDROLOGIC INFLUENCES
4. ECONOMICS

FACTORS TO CONSIDER
IN SAMPLE HANDLING, PRESERVATION AND TRANSPORTATION

1. RELATIONSHIP TO OBJECTIVES
2. SAMPLE CONTAINERS
3. PRESERVATION AND FILTRATION REQUIREMENTS
4. SAMPLE STORAGE
5. TRANSPORTATION AVAILABLE

FACTORS TO CONSIDER
IN LAB AND FIELD ANALYSES

1. RELATIONSHIP TO OBJECTIVES
2. CAPABILITIES - PARAMETERS AND VOLUME OF WORK
3. ACCURACY REQUIREMENTS OF STUDY

PROGRAM
FLEXIBILITY

1. RELATIONSHIP TO OBJECTIVES
2. CONTINUOUS FEEDBACK OF DATA
3. CONTINUOUS REVIEW

OHIO RIVER DIVISION'S
LABORATORY MASTER FILE SYSTEM

by
Henry C. Jackson¹

1. General. Before the creation of the LABMASTER system, the division water quality laboratory received samples from the field, tested them, and prepared reports of test results manually. LABMASTER was created as a means of systematically recording standard information about samples entering the laboratory, preparing bench sheets describing all samples logged into the system, recording test results, billing districts for testing, preparing reports of completed and pending results, and providing management tools for efficient operation of the laboratory. Field results were stored with laboratory results for a given sample to insure a complete information base. While the system was originally designed to run on the G437 computer, the G437 computer can no longer provide adequate service (economy, turnaround time, special plotting packages, time sharing, etc.). Therefore the information retrieval portion, as well as the field data storage portion of the system, has been moved to INFONET (Computer Sciences Corporation INFONET) and only the division laboratory portions will remain on the G437. Programs on the G437 should be run only by the division laboratory.

2. Operating Environment. The G437 portion of the system is designed to run on a Corps standard 64K G437 computer with 9 disc drives, 6 7-track tape drives, card punch, card reader, and line printer. The INFONET portion of the system is designed to run on a Computer Sciences Corporation UNIVAC 1108 computer with associated peripherals including disc drives and 9-track tape drives. Access to the INFONET system is made thru COPE-1200 terminals with card reader, card punch, line printer, and 7-track tape drives as well as thru a variety of low-speed interactive terminals.

3. Programs. The LABMASTER system is currently supported by 12 programs on the G437 for division laboratory use and 20 programs on INFONET for storage and retrieval by districts. Other programs are available but not supported on the G437. Detailed descriptions of these programs may be found in Appendix B. The programs are as follows:

G437

- a. 401H0130 Laboratory Master File Addition
- b. 401H0140 Laboratory Master File Update
- c. 401H0150 Laboratory Master File Billing
- d. 401H0170 Laboratory Master File Station List
- e. 401H0190 Laboratory Master File Bench Sheet
- f. 401H0200 Laboratory Master File Station Correction
- g. 401H0210 Laboratory Master File Table

¹ Ohio River Division, Reservoir Control Center, Water Quality Section
Paper 4

- h. 401H0220 Laboratory Master File Sample Age
- i. 401H0225 Laboratory Master File Management Information
- j. 401H0310 Laboratory Master File Station-Identification Correction
- k. 401H0430 Laboratory Master File INFONET Transfer
- l. 401H0440 Laboratory Master File Interim Bill

INFONET

- a. 401265 Laboratory Master File Sorted General Report
- b. 401340 Laboratory Master File Field Data Input
- c. 401CAT Laboratory Master File Catalog
- d. 401EDT.JA Laboratory Master File Pre-edit
- e. 401GET Laboratory Master File G437 to INFONET Transfer (Translate and Separate)
- f. 401GEX Laboratory Master File G437 to INFONET Correction Transfer
- g. 401IND Laboratory Master File Station Selection by River Index
- h. 401LOD Laboratory Master File System 2000 Data Base Load
- i. 401MER Laboratory Master File Update and Merge
- j. 401OBY Laboratory Master File Define and Equate Files
- k. 401RET Laboratory Master File Data Retrieval
- l. 401RIV Laboratory Master File River Data Report
- m. 401SEL Laboratory Master File Data Selection by Type
- n. 401SEL Laboratory Master File Data Selection
- o. 401STA Laboratory Master File Station Statistical Summary
- p. 401STD Laboratory Master File Statistics
- q. 401SUB Laboratory Master File Station-ID Selection

SYSTEM FLOWCHART FOR G437 UPDATE

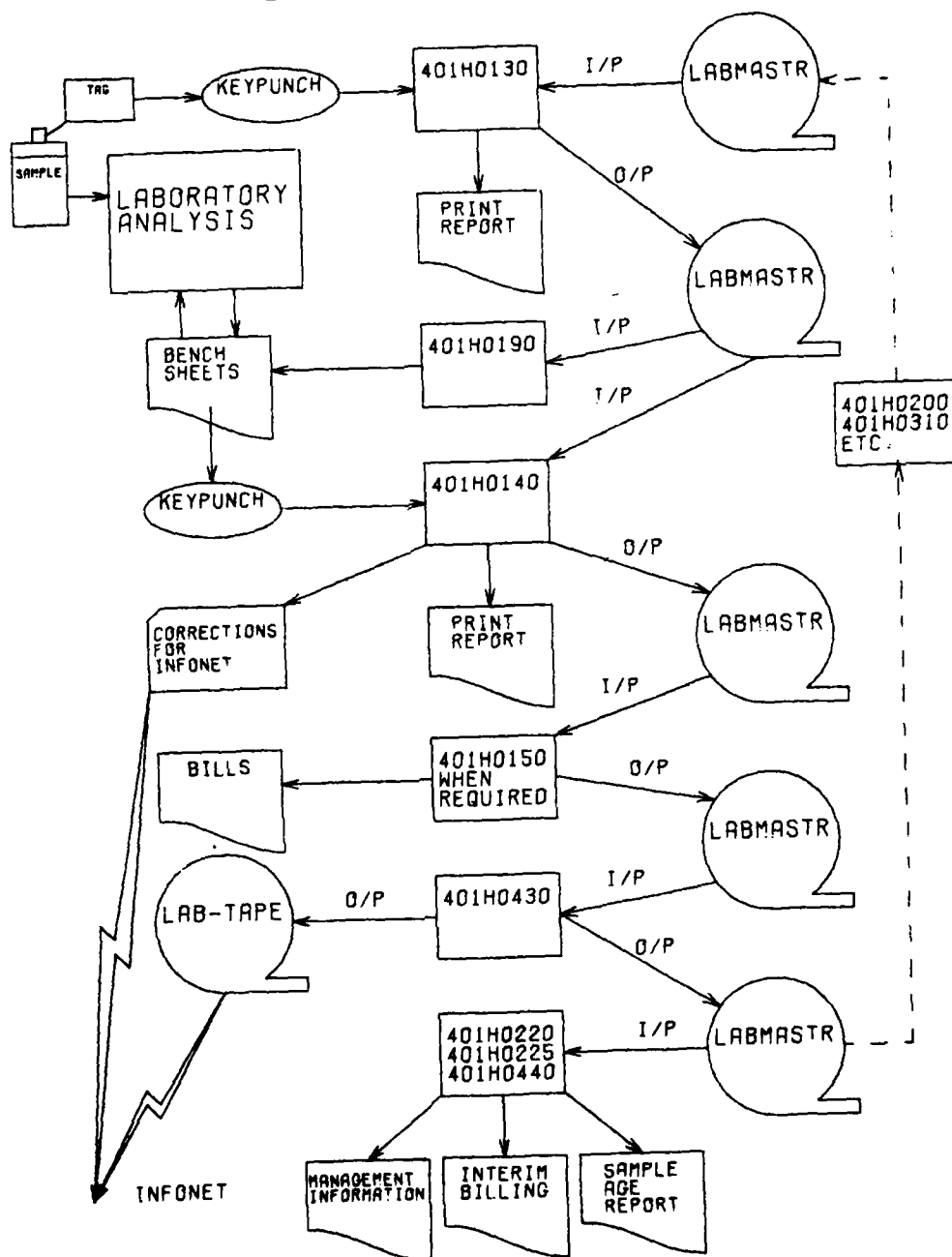


Fig. 1

- r. 401SUM Laboratory Master File Station-Day Summary
- s. 401UNP Laboratory Master File Unpunctuated File Creation
- t. 401XYZ Laboratory Master File XYZ Coordinate Append

4. Interaction of Programs on the G437. When a water quality sample is received by the division laboratory, it is assigned a laboratory number. The tag attached to the sample contains information about the sample which is keypunched directly from the tag (see figure 1 for system flowchart) run with 401H0130 to log the sample into the system. Laboratory bench sheets are prepared by running 401H0190. Management information programs such as 401H0220 or 401H0225 provide guidance to laboratory personnel in choosing which tests to perform. After tests are performed and results recorded on bench sheets, input to 401H0140 is keypunched directly from the completed bench sheets. Billing for laboratory work is accomplished by 401H0150 which is run on a quarterly basis. Interim billing information is obtained by running 401H0440 during each update cycle. Information to be transferred to INFONET is obtained from the LABMASTR tape (LABMASTR is the 8-character tape file name used on the G437) by 401H0430. A complete list of stations may be obtained by running 401H0170. Various programs such as 401H0200 or 401H0310 may be used to correct information after it has been placed on a LABMASTR tape file.

5. Interaction of Programs on INFONET. Laboratory results on tape extracted by 401H0430 from the G437 are transmitted to INFONET via the COPE-1200 terminal and placed on a tape file (4 backup tapes are maintained). Files referenced to the current data in each of the district water quality INFONET libraries are defined and equated by 401OBY (see figure 2 for system flowchart). Necessary character translation and separation by district code is done by 401GET. Special correction files in each district library are created by 401GEX on an as-needed basis. Field data input to INFONET may be accomplished by using 401340 to generate LABMASTER format data records (or by any suitable user-coded program). Disc files created by 401GET, 401GEX, 401340, etc., may be edited by 401EDT.JA before placing the data on a tape master file (LABFIL) with 401MER (see figure 3 for system flowchart). A detailed catalog of a LABFIL may be obtained by running 401CAT. A disc-resident subset of the data on a LABFIL may be obtained by running 401SEL, 401SUB, or 401RET. A formatted report of information on a LABFIL or a subfile may be obtained by running 401265, 401RIV, or 401SUM. A system 2000 data base loader input may be obtained by running 401LOD. Standard statistical analysis of the information in a sorted subfile may be obtained by running 401STD or 401STA. Temporary subfiles (one file per project) for binding into a quick-access unpunctuated tape may be obtained by running 401UNP. Raw data records may be modified for plotting purposes by running 401XYZ which appends the river mile, elevation, and distance from left bank to each record. Stations may be selected by 401IND which allows the user to specify STORET stream indices and river miles as selection criteria. Subsequent raw data retrieval may be made by 401RET. A detailed description of each station is contained in a special random access file called DESFIL which may be accessed by any program. Examples of program use may be found in Appendix C.

SYSTEM FLOWCHART FOR DATA TRANSFER

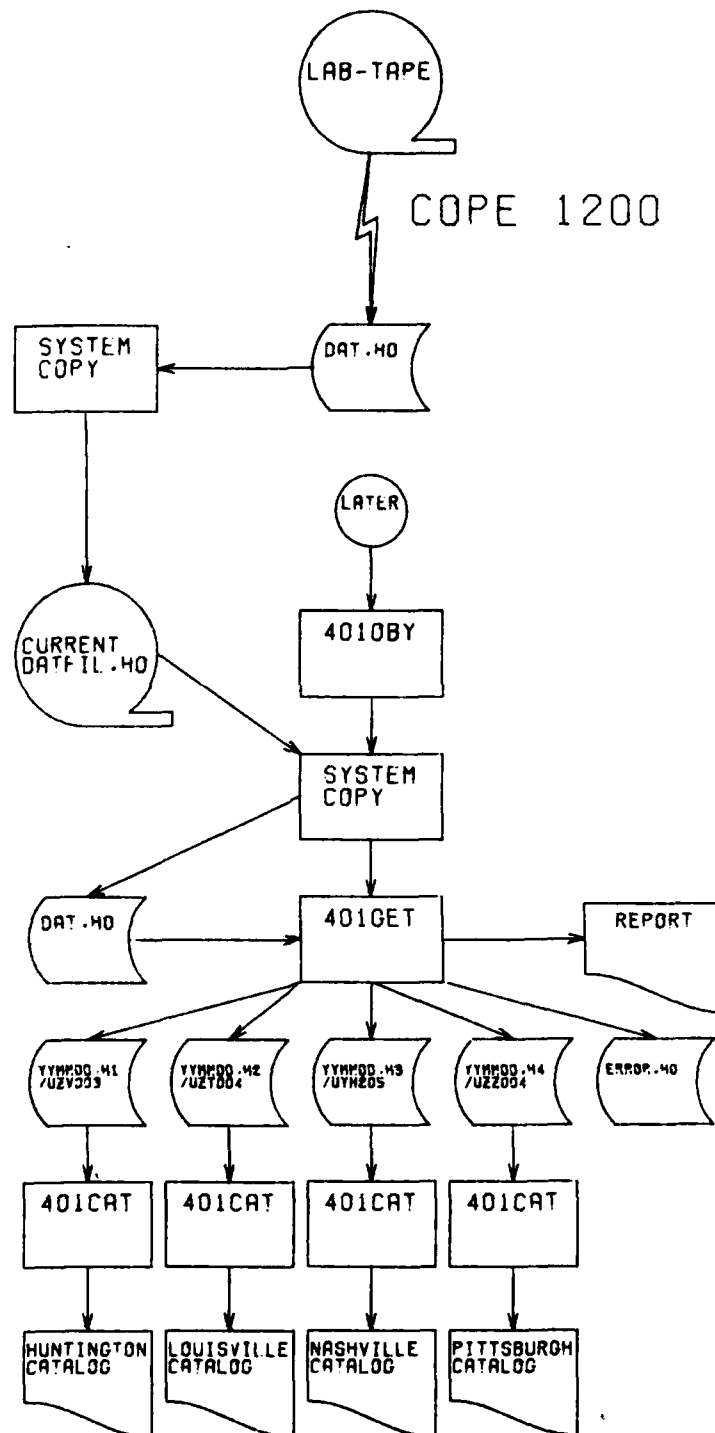
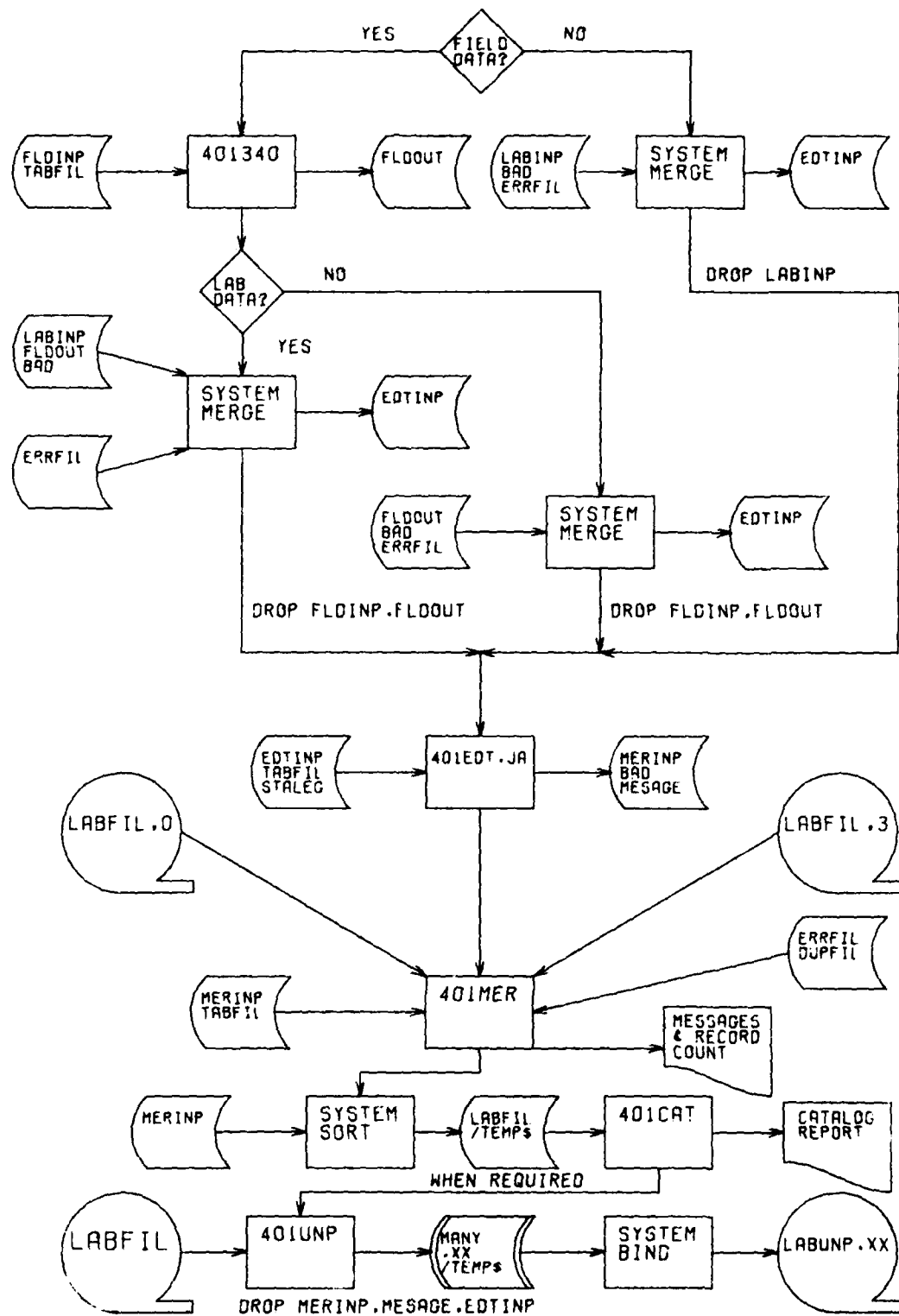


Fig. 2

SYSTEM FLOWCHART FOR INFONET UPDATE



BAD AND ERRFIL SHOULD BE CORRECTED WITH SYSTEM EDITOR BEFORE NEXT UPDATE

6. Data Elements. A standard method of describing the test, location, and time of sampling was required. The following descriptive terms have been adopted:

| | |
|----------------------|--|
| TEST | Five character STORET parameter code for a standard laboratory test (see Appendix A for list). |
| LABORATORY NUMBER | Five-digit number assigned to each sample by the laboratory. |
| STATION | Nine-character location code described below: |
| DISTRICT | One-digit code |
| 1 | Huntington District |
| 2 | Louisville District |
| 3 | Nashville District |
| 4 | Pittsburgh District |
| 5-9 | Other |
| PROJECT | Three-character code assigned by district |
| TYPE | One digit type of sampling code assigned by district |
| 1 | River |
| 2 | Lake |
| 3 | Benthic |
| 4 | Sewage |
| 5 | Industrial |
| 6 | Potable Water |
| 7 | Subsurface Water |
| 8 | Surface Soil |
| 0 or 9 | Other |
| RIVER MILE | Four-digit code assigned by district |
| IDENTIFICA- TION | Thirteen-digit time and depth code described below. All fields are right-justified and zero-filled: |
| YEAR | Two-digit calendar year |

| | |
|--------------------|---|
| MONTH | Two-digit month |
| DAY | Two-digit day of month |
| TIME | Four-digit time of day (24-hour clock) |
| DEPTH | Three-digit depth of sampling |
| CONCENTRATION | Seven-character test result described below: |
| TRACE INDICATOR | |
| L | Minimal test result |
| G | Maximal test result |
| N | Normal |
| VALUE | Six-character numeric test result (decimal point, plus sign, and minus sign are permitted). |
| COST | Five-digit test cost in cents |
| BILL INDICATOR | "*" if billed; space otherwise |
| EPA INDICATOR | Three-character flag described below: |
| EPA | Data on record previously transferred to STORET |
| REV | Test result on record has been changed and not transferred to STORET |
| ANY | A normally described Station-ID appears on this record |
| NNN | Data not transferred to STORET |
| INFONET INDICATOR | One-character indicator (with meaning only on G437 system) as described below: |
| I | Data transferred to INFONET |
| N | Data not transferred to INFONET |
| TRANSFER INDICATOR | One-character indicator (with meaning only on G437 system) as described below: |
| T | Data transferred to district LABMASTER |
| N | Data not transferred to district LABMASTER (obsolete) |

A special coding system has been devised to facilitate the naming of main stem Ohio River stations. This coding system is as follows:

| <u>POSITION</u> | <u>INFORMATION</u> |
|-----------------|---|
| 1 | District code |
| 2-3 | OR |
| 4 | Hundreds position of river mile (i.e., 6 for river mile <u>625.7</u> from <u>Pittsburgh</u>) |
| 5 | TYPE CODE (usually 1) |
| 6-8 | Tens, units, and tenths positions of river mile (i.e., 257 for river mile <u>625.7</u>) |
| 9 | Relative position in tenths of cross section (0 for left bank to 9 for right bank) |

7. File Names and Structure of the G437 System. Tape and disc files on the G437 system must be eight-character names. The master tape file containing all laboratory results as well as work-in-progress is called LABMASTR. The LABMASTR tape has a standard GE tape label with a VALUE-OF-ID of "LABMASTR." Records are 64 characters long and blocks are 100 records long with block serial numbers. The first record of the tape consists of all zeros except columns 33-40 which contain the last billing date. The remaining data, as described below, are sorted in ascending order on the first 32 columns:

| <u>POSITION</u> | <u>INFORMATION</u> |
|-----------------|--------------------|
| 1-5 | TEST |
| 6-10 | LABORATORY NUMBER |
| 11-19 | STATION |
| 20-32 | IDENTIFICATION |
| 33-39 | CONCENTRATION |
| 40-44 | COST |
| 45 | BILL INDICATOR |
| 46-59 | Not Used |
| 60-62 | EPA INDICATOR |
| 63 | INFONET INDICATOR |
| 64 | TRANSFER INDICATOR |

The table of acceptable STORET parameters and associated information is stored in a permanent disc file called "XXXXXXXX" on the "IDSEDIT1" disc. Records are 64 characters long and the blocking factor is 60. Data are sorted in ascending order on the first 5 columns. The record description is as follows:

| <u>COLUMN</u> | <u>INFORMATION</u> |
|---------------|--|
| 1-5 | STORET Parameter Code |
| 6-13 | Line 1 of Title |
| 14-21 | Line 2 of Title |
| 22-29 | Line 3 of Title |
| 30-34 | Cost in Cents |
| 35-42 | Obsolete Number in 401H0160 |
| 43-46 | Shelf Life Code |
| 47 | Sample Type Code 1=WATER 3=MUD, otherwise blank |
| 48 | Minimum Value Flag (L if trace values are reported; space otherwise) |
| 49-54 | Minimum Acceptable Value |
| 55 | Maximal Value Flag (G if maximal values are reported; space otherwise) |
| 56-61 | Maximum Acceptable Value |

The code which is used in the transmission of information from the G437 to INFONET is called "L*B-TAPE." It is an unlabeled card image tape with 80 character records unblocked. The first 64 characters of each record are the same as on the LABMASTR tape and the remaining 16 characters are blank.

8. File Names and Structure on INFONET. Tape and disc files on INFONET have six-character names (may be less than 6) and an optional two-character version which may be appended to the end of the file name (separated by a period). Special tape files defined as "generation data sets" and providing backup capability ("grandfather," "father," son," etc.) are described by affixing a comma and a backup number to the end of a tape file name (i.e., MYTAPE.H1,2). File names may be further qualified by appending a slash and the library name (i.e., 401MER/ORDLAB would be in the shared division library and ERRFIL.H1/UZV003 would be in

the Huntington District library). Unless otherwise stated, all files have standard keys and records are 64 characters long. The standard sort sequence is ascending order of columns 11-32 and 1-5. Standard version names for LABMASTER files are as follows:

.H0 for Division Office
.H1 for Huntington District
.H2 for Louisville District
.H3 for Nashville District
.H4 for Pittsburgh District

Standard file names (with version represented by xx) are as follows:

| | |
|-------------|---|
| BAD.xx | File containing records rejected by editor program (401EDT.JA/ORDLAB) to be corrected with the system editor and re-edited. |
| DAT.H0 | Temporary disc file containing laboratory data most recently transferred to INFONET. |
| DATFIL.H0,0 | Tape file containing laboratory data transferred to INFONET (,0 indicates most recent data). |
| DESFIL.xx | File containing station description information in the EPA STORET format as described in the description for program 401DES, DESFIL may be accessed randomly or sequentially because it has 3-word nonstandard keys corresponding to the station code and STORET record number. |
| DUPFIL.xx | File containing duplicate records detected by master file update program (401MER/ORDLAB) to be checked and (in most cases) deleted with the system editor. |
| EDTINP.xx | File containing input data for editor program (401EDT.JA/ORDLAB). |
| ERRFIL.xx | File containing records rejected by master file update program (401MER/ORDLAB) to be corrected with the system editor and re-edited. |
| FLDINP.xx | File containing field data input to be run with field data format program (401340/ORDLAB). |
| FLDOUT.xx | File containing field data after reformatting into standard format by 401340/ORDLAB. |
| LABCOP.xx | Tape resident duplicate of laboratory master file to be used as backup. |

| | |
|---------------|---|
| LABDEF | System 2000 data base definition file. |
| LABLOD | System 2000 data base load instruction file. |
| LABFIL.xx | Tape resident input laboratory master file containing all district laboratory results. |
| LABFIL.xx,3 | Tape resident output laboratory master file (3 indicates that 3 generations of backup tapes are to be maintained) which is created by 401MER/ORDLAB. |
| LABINP.xx | File containing laboratory data after merging all laboratory provided files (i.e., 770105.xx, 77999.xx, etc.). |
| LABUNP.xx-UNP | Unpunctuated tape resident file containing one subfile for each project (input files are generated by 401UNP/ORDLAB) which is used for quick access of data for an entire project. |
| LOADDO | System 3000 data base load "do file." |
| MERINP.xx | File containing all data to be updated on master file by 401MER/ORDLAB. |
| STAFIL | File containing stations selected by 401IND which will later be used as input to 401RET. |
| STALEG | File containing acceptable stations for use by 401EDT.JA/ORDLAB. Station number is in positions 1-9, maximum depth is in positions 10-12, and EPA transfer code is in positions 13-15 ("EPA" or blank). This file is sorted on positions 1-9. |
| TABFIL | File containing information on every acceptable STORET parameter (same format as XXXXXXXX on IDSEDT1 on G437). |
| yyymmdd.xx | File containing laboratory information placed on INFONET on year yy, month mm, and day dd (i.e., 770109.xx would contain information for 9 Jan 77). |
| yy99zz.xx | File containing corrections to laboratory information on INFONET (i.e., 779994.xx) which should be carefully inspected by district personnel before using in an update. |

9. Special INFONET System Routines. There are a few system routines on INFONET which are especially useful in massaging and organizing data files. They are well documented in the INFONET manuals. The following is a list of some of the more useful available routines:

a. System Editor. This is a powerful text editor which is invoked by the command !EDIT file name. The user should refer to the manuals or pocket guide for a full list of editor commands.

b. System Sort. This is an easy to use sorting routine which may be executed by a single command. A typical sort command to sort a file in the standard LABMASTER sequence would be !SORT 64 inputfile outputfile 11,22 1,5.

c. System Merge. This program provides an easy way to concatenate two or more files into another file without disturbing their order, or in the case of already sorted files, it may be used to create a single sorted file from two or more input files. A typical merge command used with LABMASTER data files is !MERGE,COPY 64 file1,file2,file3 outputfile.

d. SYSTEM 2000. System 2000 is an extremely powerful data management language which can enable a user to perform a limitless variety of queries on a properly created database. It is essential that the user have a System 2000 manual before attempting to use this program.

e. ALADIN and REPORT II. ALADIN is a more economical data management language which provides an interface with the REPORT II report writer which enables the user to prepare one-time especially formatted reports quickly and easily.

STATISTICAL INTERPRETATION OF WATER QUALITY DATA

BY

James L. Grant ¹

INTRODUCTION

Water quality data represents information about the state and behavior of the chemical and physical aspects of a hydrologic system. Analyses of this information are directed toward separating the consequential from the inconsequential aspects of the data, and point to a better understanding of the underlying causal and random mechanisms of the system.

Uncertainty in the water quality data may arise from a multitude of sources, including unknown or uncertain system inputs, uncertain system responses to known inputs, and observational errors resulting from sampling, storage and testing procedures. In spite of these uncertainties, knowledge often is available which allows a deterministic appraisal of at least a portion of the response of the system being monitored, so that the system response may be partitioned into a deterministic portion and a residual. In this discussion, techniques are presented which enable a rational analysis of the observed data to be made with a minimum a priori knowledge of the behavior of the system. The procedures are designed to aid in the identification and characterization of both the random and non-random components of the data.

Once these components of system response have been identified, causative mechanisms for the observed non-random behavior may be sought, and the non-random portion of the data used for calibration of the deterministic models thus developed. The statistical properties of the random components of the data may be used in conjunction with the deterministic system response to predict average and extreme system behavior.

The statistical analysis of data is an integral part of the development and calibration of models of hydrologic systems. Further, because the process described above allows the observed system behavior to be separated into deterministic and "random" components, such analyses generally should allow realistic estimates to be made of "normal" and "extreme" conditions which may occur in the future.

FITTING OF DETERMINISTIC MODELS TO SAMPLE DATA

In order to provide motivation for the study of the abstract concepts of spectral analysis of data to be presented later in this section, we begin our discussion of data analysis by examining the fitting of deterministic models to observed water quality data.

¹Nuclear Engineering Company, Inc., Louisville, Kentucky. Formerly Chief Hydrologist, Law Engineering Testing Company, Marietta, Georgia.

Such analyses have important practical applications. For example, water quality models are often calibrated by adjusting the model parameters until the model output is "close" to the observed monitoring data. The adequacy of agreement between observed data and model output may be determined subjectively or objectively.

The most familiar objective criterion is the method of least squares. In this method, the parameters of the model are adjusted until the sum of the squares of the differences, between predicted and observed values is a minimum. In many cases, the differences, or residuals, obtained by this procedure have desirable statistical properties, and thus the prediction of the likelihood of occurrence of extreme events through such an analysis is enhanced.

As an example of the techniques of parameter estimation by the method of least squares, suppose we generate a sequence of temperature measurements over a one year period in a standard stream. We assume that the (unknown) function relating temperature to time is:

$$T = 68 + 5 \sin (2\pi t) + e$$

Where t is time in years, and e is a normally - distributed random variate from $N(0, 16)$. Using the above equation and a table of normal random variates, the following "observations" may be generated.

| t | $T' = T - e$ | e | T |
|------|--------------|-------|-------|
| 0 | 68 | -0.12 | 67.88 |
| .25 | 73 | 2.36 | 75.36 |
| .5 | 68 | 1.65 | 69.65 |
| 0.75 | 63 | -3.44 | 59.56 |
| 1 | 68 | -1.68 | 66.32 |

In a practical problem, we will have available the observations T , and the parametric model

$$T' = a + b \sin (2\pi t).$$

Our objective is to use the observed data to estimate the parameters a and b as well as the population parameters of the distribution of the random component of the observations. The error associated with the i th observation is defined by the relation

$$e_i = T_i - [a + b \sin (2\pi t)]$$

Using the usual procedures of least squares analysis, we obtain the parameter estimates

$$\hat{a} = 67.54$$

and

$$\hat{b} = 7.9.$$

An analysis of the residuals e_i yields the following estimates of the sample mean and variances

$$\bar{e} = 0$$

and

$$S_e^2 = 5.84.$$

Using our derived relation, we may estimate the temperature at $t = 0.5$ which, on the average, will be exceeded only five percent of the time. To do this, we first compute the deterministic portion of the system response

$$\text{from the relation } T' = \hat{a} + \hat{b} \sin(2\pi t) = 67.754$$

we then compute $e_{0.95}$

$$e_{0.95} = \bar{e} + Se(Z_{0.95}) = 3.981$$

where Z is the .95 percentile point of the standard normal distribution.

$$T_{0.95} = 67.54 + 3.981 = 71.735$$

Since in this example we know the model from which the observations were generated, we may compute the actual, or true, value of $T_{0.95}$ as

$$T_{0.95} = 68 + 4(1.645) = 74.58$$

We see by comparison of the above estimate of $T_{0.95}$ with the true value that our

best estimate of $T_{0.95}$ is too low by almost 3%. For some applications, the risk of making such an error may not be acceptable, and in these instances we are led to the consideration of tolerance limits of our estimate. By a tolerance limit we mean an estimate of (for example) $T_{0.95}$ for which we can be assured that if we make such an estimate in a large number of cases, the inequality

$$T_{0.95}(\text{estimated}) \geq T_{0.95}(\text{true})$$

will be true at least a specified number of times. To see how such an estimate may be constructed, we first address the problem of describing confidence limits on our estimates of the parameters a and b . A confidence limit in this case means a region in the $a - b$ plane in which, if we construct such regions for a large number of cases, the true values of a and b will lie within at least a specified percentage of the regions. Such a region may be constructed as follows (Grant, 1973): Define the regression component of the errors as

$$\text{Reg}(e) = \sum_{i=1}^N [\hat{a} + \hat{b} \sin(2\pi t_i) - \bar{T}]^2$$

and the residue of the errors as

$$\text{Res (e)} = \sum e_i^2 - \text{Reg (e)}$$

The function Reg (e) has rank 2 and Res (e) has rank N-2 degrees of freedom. The α percent confidence region for a and b is then defined as the region of all pairs of points (a,b) which satisfy the inequality

$$\frac{\text{Reg(e)}}{\text{Reg(e)} - N-2} \leq \frac{2}{100} F \left(\frac{\alpha}{2}, N-2 \right)$$

Figure 1 shows a sketch of the 95% confidence region for a and b in our example.

Once we have available the 95% confidence region for the parameters a and b, we may define the 95% tolerance limit of the estimate T' as the maximum of

$$T' = a + b \sin (2\pi t)$$

over the 95% confidence region. The population parameters of the residual errors (in this case, the standard deviation), are estimated from the errors evaluated at the critical values of a and b. For our example, we find by trial and error that

$$T'_{0.95} = 71.51$$

and thus,

$$S_e = 6.62$$

Our 95% tolerance limit is thus

$$\begin{aligned} T_{0.95,0.95} &= T'_{0.95} + S_e (Z_{0.95}) \\ &= 81.40 \end{aligned}$$

The above results have been presented for the case of linear least squares, that is, for the case for which the unknown parameters a and b appear linearly in the postulated model. The results can be generalized in a natural manner to include non-linear cases. The details of this extension are presented in Hartley, 1964; Grant, 1973; and Wallace and Grant, 1971.

DATA ANALYSIS WHEN THE UNDERLYING DETERMINISTIC FACTORS ARE NOT KNOWN

In the preceeding section, we examined an application of the method of least squares in the analysis of data. In that example, we assumed that the form of the underlying deterministic mechanism was known. This assumption allowed the analysis to be performed by a straightforward separation of the observed responses into a deterministic and random components. The statistical properties of the parameter estimates and the residual errors then allowed the computation of statistical estimates of the system response.

In this section, we assume that the form of the deterministic mechanism is

unknown, and we examine techniques of analysis which allow us to infer certain properties of the deterministic mechanisms as well as the stochastic properties of the system. From these inferences, we then would postulate forms of these processes and proceed with our analysis as before.

Whenever we are attempting to analyze a set of data, we generally find that the data set has two properties which constrain our techniques of analysis. First, although the data may be representative of a continuous system, we will have knowledge of the state of the system only at discrete points. Second, we will find that our sequence of data covers only a finite time period, so that we are unaware of the state of the system except during this interval. Given these constraints, we must first decide whether to treat the system being monitored as a discrete system or, by interpolating between the given data points, to attempt to approximate the continuous nature of the system. The restriction imposed by the finite length of our sample requires us to treat our system as periodic with a period equal to our record length, or to extend the record based upon known or assumed properties of the system.

In this discussion, we will confine ourselves to an examination of techniques of the spectral analysis of discrete periodic systems. Extensions of these techniques to continuous aperiodic systems are presented in Kisiel, 1969, and to discrete aperiodic systems in Kisiel, 1969, and Oppenheim and Schafer, 1975.

Suppose we have a sequence of N observations x_n , $n = 0, 1, \dots, N-1$. We may write the Fourier transform of this sequence (Oppenheim and Schafer, 1975) as follows:

$$X(k) = \sum_{n=0}^{N-1} x_n \text{EXP} \left[-j \frac{(2\pi)nk}{N} \right]$$

Where the $X(k)$ are the Fourier coefficients of the series representation of the sequence x_n , that is

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \text{EXP} \left[j \frac{(2\pi)nk}{N} \right]$$

Statistical properties of the sequence x_n which are of interest in data analysis include the mean, the variance, and the autocorrelation and autocovariance functions. These quantities are defined as follows:

$$\begin{aligned} \text{Mean} \quad \bar{x}_n &= E(x_n) \text{ for each } n, \\ \text{Variance} \quad \sigma_n^2 &= E(x_n - \bar{x}_n)^2 \text{ for each } n, \end{aligned}$$

Autocorrelation sequence

$$\phi_{xx}(n, m) = E(x_n \cdot x_m),$$

Autocovariance sequence

$$\gamma_{xx}(n, m) = E[(x_n - \bar{x}_n) \cdot (x_m - \bar{x}_m)] = \phi_{xx}(n, m) - \bar{x}_n \bar{x}_m.$$

For a stationary process, we may write, since means and variances are independent of time,

$$\begin{aligned}\bar{x} &= E(x_n) \\ \sigma_x^2 &= E(x_n - \bar{x})^2 \\ \phi_{xx}(m) &= E(x_n \cdot x_{n+m}) \\ \gamma_{xx}(m) &= \phi_{xx}(m) - \bar{x}^2\end{aligned}$$

For ergodic sequences, we may estimate the above averages as the time averages of the single sequence. Various formulas have been proposed to compute estimates of the averages of a data sequence. For example, means and variances may be computed by the formulas:

$$\begin{aligned}\bar{x} &= \frac{1}{N} \sum_{i=0}^{N-1} x_i \\ S_x^2 &= \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \bar{x})^2\end{aligned}$$

The autocovariance sequence may be approximated by the relation

$$k_x(k) = (N-k)^{-1} \sum_{i=0}^{N-k} (x_i - \bar{x})(x_{i+k} - \bar{x})$$

and the autocorrelation function by the relation

$$K_x(k) = k_x(k) + \bar{x}^2.$$

The spectral density function of a sample is defined as the Fourier transform of the autocovariance sequence. An unsmoothed estimate of the one-sided spectral density function is given by

$$\tilde{G}_x(f) = \frac{2}{M} \left[K_x(0) + 2 \sum_{r=1}^{M-1} K_x(r) \cos\left(\frac{\pi r f}{f_c}\right) + K_x(M) \cos\left(\frac{\pi M f}{f_c}\right) \right]$$

Where $1/f_c$ equals twice the sampling frequency. This critical frequency, known as the Nyquist frequency, determines the behavior of the analysis. Frequencies higher than the critical frequency are not visible to the analysis; however, the variance associated with these frequencies may be reflected, or "aliased" into the visible spectrum.

The above estimate of the spectral density function is not particularly suitable for analysis because it is unstable with respect to sample size. Various techniques of obtaining a smoothed, or stable, estimate are presented in Oppenheim and Schaffer, 1975, and a rather simple procedure is presented in Kisiel, 1969.

The primary purpose of spectral analysis of monitoring data is to allow some separation of the mechanisms of variance in the data, and to simplify

the statistical properties of the data. When appropriate, and properly used, spectral analysis can be a powerful tool to aid in the interpretation of data. The use of spectral analysis can be illustrated by the following simple example. Suppose we sample a system generated by the process

$$x(t) = 5 \sin (2\pi t)$$

at intervals of 0.25. We obtain the sequence 0,5,0,-5,---. Suppose that we have a sequence of 13 such data points which we use to estimate the spectral density function of the system. By the above equations, we find

$$K_0 = 11.54$$

$$K_1 = 0$$

$$K_2 = -11.32$$

$$K_3 = 0$$

$$K_4 = 11.11$$

and

$$\tilde{G}(f) = 0.5[11.54 - 22.72 \cos \pi f + 11.11 \cos 2\pi f].$$

The spectral density function is to be evaluated at the discrete frequencies $f = 0.5k$, and we find

$$\tilde{G}(0) = 0.04$$

$$\tilde{G}(0.5) = 0.22$$

$$\tilde{G}(1) = 22.69$$

$$\tilde{G}(1.5) = 0.22$$

$$\tilde{G}(2) = -0.04.$$

A plot of the variance spectra is shown in Figure 2. The preponderance of the variance associated with $f = 1$ indicates that our data sample was obtained from a periodic process with period equal to one year. Of course, spectral density functions of real data usually are more complicated than the above example, and the interpretations of the variance spectra is not so straightforward. Examples of interpretation of more complicated spectra are presented in Kisiel, 1969.

Once the deterministic contributions of a data sequence are identified, the contributions of these deterministic components can be removed from the data. The residuals obtained by this process represent the random components of the system, and may be analyzed by standard statistical methods. The results of such analysis allow the synthesis of the physical system and a prediction of extreme conditions which may occur within the system in the future.

Paper 5

COMMENTS ON SAMPLING FREQUENCIES

The above discussion illustrates the requirements of a rational water quality sampling program. The frequency of sampling must satisfy the Nyquist relation for the highest frequency which the subsequent data analysis will consider. Because natural systems rarely are governed by a pure harmonic process, sampling at a multiple of two or three times the Nyquist frequency is recommended (Kiesel, 1969).

Frequencies too high to be visible to the analysis of a data set either will be folded back into the visible frequency range, a phenomenon known as aliasing, or will appear as random noise to the analysis. Figure 3 shows a conceptual diagram of the aliasing of a high frequency process as a result of widely-spaced sampling.

Requirements upon sample size and the length of the autocovariance sequence are dictated by the statistical properties of the estimates of the averages and the spectral density function of the data. The maximum number of serial lags, m , used in computations generally should be 10% of the sample size or less. For a specified standard error of estimate e of the spectral density function, the sample size required is $N = m/e^2$, and the length of record required is $N\Delta T$ (Kiesel, 1969).

SUMMARY

Statistical analysis of water quality data usually is performed in order to utilize the information contained within the data set to obtain an understanding of the system from which the data were obtained. Natural systems generally exhibit characteristics which are in part deterministic and in part random. A rational analysis of water quality data requires that the deterministic elements of the system be identified and their influences deleted from the data. The stochastic and random components of the system then can be characterized statistically.

Techniques for the analysis of data exhibiting deterministic and stochastic properties have been discussed, and simple examples of the use of the methods have been presented. The implications of the analysis techniques presented are discussed in terms of the requirements which these techniques impose upon data collection programs.

REFERENCES

- Grant, J. L.: "Statistical Frequency Analysis by Optimization of Density Functions to Histograms", Thesis presented to the Faculty of the Division of Graduate Studies and Research, Georgia Institute of Technology, at Atlanta, Georgia Institute of Technology, at Atlanta, Georgia, in 1973, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.
- Hartley, H.O.: "Exact Confidence Regions for the Parameters in Non-Linear Regression Laws", Biometrika, Vol. 51. University College, London, G.B., 1964, pp. 347 - 353.
- Kisiel, C.C.: "Time Series Analysis of Hydrologic Data", Advances in Hydroscience, V.T. Chow, ed., Vol. 5, Academic Press, New York, N.Y. 1969. pp. 1 - 119.
- Oppenheim, A. V., and Schaffer, R.W.: Digital Signal Processing, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975.
- Wallace, J. R., and Grant, J.L.: "A Least Squares Method for Computing Statistical Tolerance Limits", Water Resources Research, Vol. 13, No. 5, Oct., 1977; pp. 819 - 823.

Figure 1
95% CONFIDENCE REGION
FOR LINEAR REGRESSION MODEL

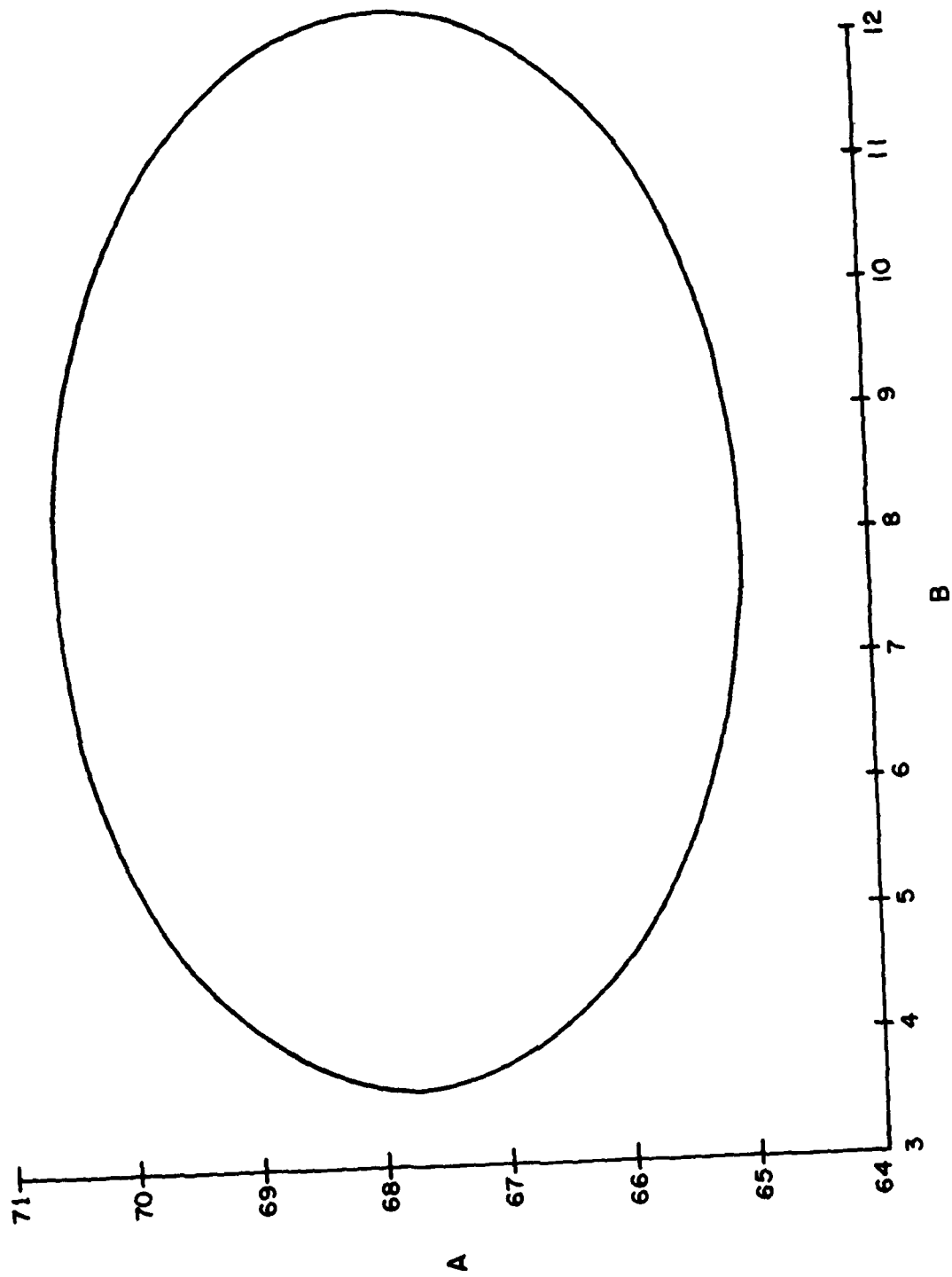
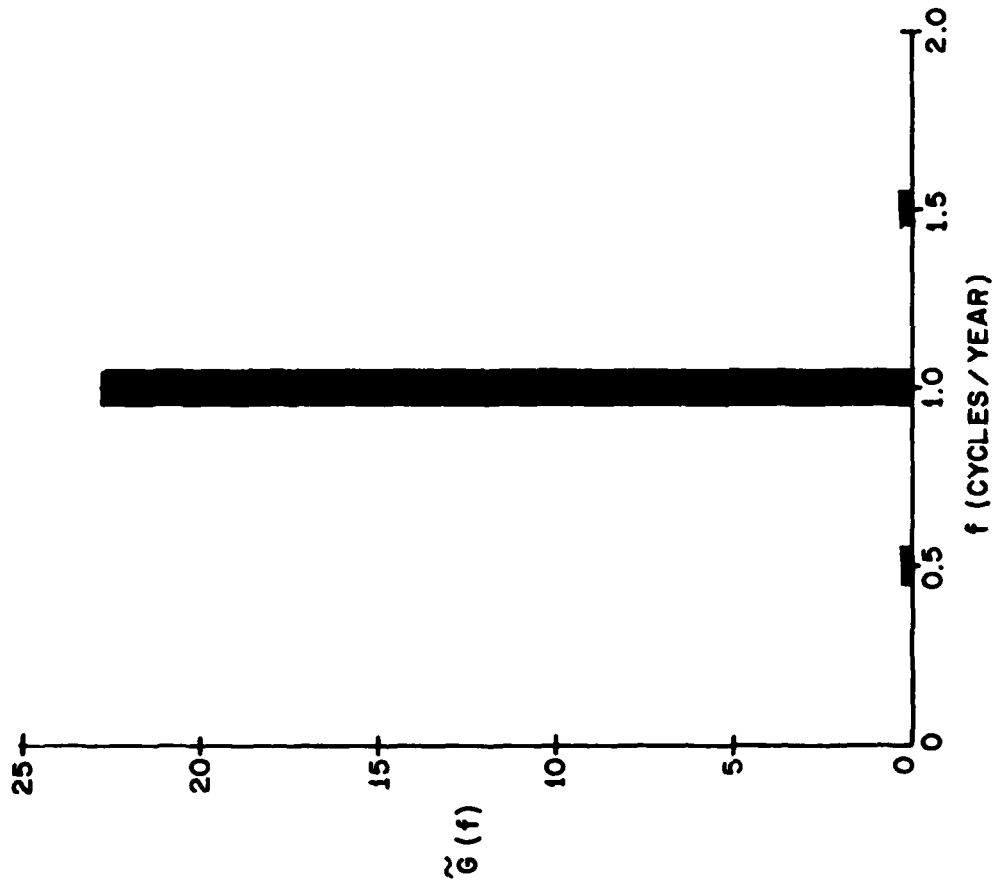


Figure 2
SPECTRAL DENSITY FUNCTION
FOR
SINUSOIDAL SIGNAL





LEGEND
 X SAMPLE
 — SIGNAL
 - - - ALIASED SIGNAL

Figure 3
 ALIASING OF
 HIGH FREQUENCY SIGNALS

DATA VALIDATION

By
James D. Ashworth¹

Introduction

This paper is a general presentation of the topic of environmental data validation. It is subdivided into theoretical considerations, methodology, and procedures presently in use. Neither computer software nor hardware are discussed. The reader is advised that terms used are defined in the Appendix.

Theoretical Considerations

Intuitively, data validation must rely upon the organization of records into groups based upon common attributes and independent components. The validating process then requires the ability to discern known intra-group properties and relationships. Group boundaries are not crossed during any test.

In general, records resulting from environmental monitoring contain five independent components; i.e., parameter name, time, and 3 dimensional location. Consequently, the theoretical number of group types or classes is $2^5 = 32$. Figure 1. is a summarization of the theoretical grouping and testing process. Each row represents an organization for all of the data to be validated.

Confining this discussion to non-redundant records, the Class I groups contain one record each. Intra-record tests of syntax, component logic, and reference comparison are possible. The Class II groups contain a complete sample permitting tests for balances and inter-parameter relationships. Class III groups can be chosen by varying any one of the four space-time components while holding the other three as well as parameter name constant. This class enables testing for continuity. Classes I thru III constitute a comprehensive validating effort. Higher classes are confounded by multiple variables rendering relationships difficult to define and cumbersome for validating purposes. In fact these data groupings are used in sophisticated modeling efforts, a major purpose for the creation of the data base.

Testing is normally performed in class order with concurrent editing. No edit can be made without a manual effort involving a multiple path decision. Ultimately, substantial evidence of a mistake must be found or exceptions must be accepted. Class I exceptions are most easily corrected since testing concentrates on the independent portion of the record. All higher classes rely on value testing. While Classes I and II exceptions tend to be distributed randomly throughout data with

respect to all components, Class III exceptions tend to be outliers. Consequently, incorrect Class III editing decisions can bias the data base. Finally, the value range test which is often associated with Class I is legitimately a part of classes III or higher inasmuch as the range is determined by experience.

Methodology

Tests of record syntax search for illegal characters based on the allowable syntax. Examples of component logic include determination that recorded year, month, day, hour, and minute do not exceed current year, twelve, the maximum for the month, 23, and 59, respectively. Reference comparisons can be made for station number, date, time, depth, parameter, and value. Within Class I, however, the value range is based on physical constraints and not experience.

Intra-sample balances for charge and total and dissolved constituents are appropriate Class II tests as well as alkalinity-acidity-pH, specific conductance-ionic strength, and dissolved oxygen-temperature relationships.

Continuity can be detected computationally or graphically. Such testing is based upon the assumption that nature rarely produces a discontinuous function. The major prerequisite is high data density with respect to the fixed independent components. A common application is the plot of depth versus concentration holding horizontal location and time constant. Computational approaches rely upon changes in the slope of the chosen function. It should be stressed that since Class III tests seek outliers, incorrect application can bias the data base.

Procedures In Use

The validating procedures used in the maintenance of four large data bases are outlined in Figure 2. It is seen that data collected with the aid of robot monitors receives somewhat different attention than that collected by manual methods. This is primarily the result of three factors, higher data but lower parameter densities and decreased opportunity for transcription errors, where data is collected mechanically. Consequently, the Geological Survey often tests for continuity of Class II relationships corresponding to line 10 of Figure 1 at monitored stations, but for Ionic Balance, Class II, of grab samples assumed to measure all constituents.

The proposed validating system of the National Climatic Center will use asymptotic singular decomposition of a matrix to streamline Classes II and III relational testing as well as provide data compaction. Value range limits based on physical constraints are contrasted with those based on experience by testing procedures for records from Orsanco

THEORETICAL MATRIX OF RECORD GROUPING CHOICES FOR DATA
VALIDATION TESTING

| INDEPENDENT COMPONENTS | | | | | | | TEST | TYPE |
|------------------------|-----|---|---|---|------|--------------|---|------|
| CLASS | PAR | X | Y | Z | TIME | | | |
| 1. I | c | c | c | c | c | Intra-Record | Syntax, Component Logic, Reference Comparison | |
| 2. II | v | c | c | c | c | Intra-Sample | Inter-Parameter Relationships, Balances | |
| 3. III | c | v | c | c | c | Inter-Sample | Continuity | |
| 4. III | c | c | v | c | c | " | " | |
| 5. III | c | c | c | v | c | " | " | |
| 6. III | c | c | c | c | v | " | " | |
| 7. | v | v | c | c | c | | | |
| 8. | v | c | v | c | c | | | |
| 9. | v | c | c | v | c | | | |
| 10. | v | c | c | c | v | | | |
| 11. | c | v | v | c | c | | | |
| 12. | c | v | c | v | c | | | |
| 13. | c | v | c | c | v | | | |
| 14. | c | c | v | v | c | | | |
| 15. | c | c | v | c | v | | | |
| 16. | c | c | c | v | v | | | |
| 17. | v | v | v | c | c | | | |
| 18. | v | v | c | v | c | | | |
| 19. | v | v | c | c | v | | | |
| 20. | v | c | v | v | c | | | |
| 21. | v | c | v | c | v | | | |
| 22. | v | c | c | v | v | | | |
| 23. | c | v | v | v | c | | | |
| 24. | c | v | v | c | v | | | |
| 25. | c | v | c | v | v | | | |
| 26. | c | c | v | v | v | | | |
| 27. | v | v | v | v | c | | | |
| 28. | v | v | v | c | v | | | |
| 29. | v | v | c | v | v | | | |
| 30. | v | c | v | v | v | | | |
| 31. | c | v | v | v | v | | | |
| 32. | v | v | v | v | v | | | |

c = Constant
v = Variable

INDEPENDENT DEPENDENT
RECORD = (PARAMETER)(X, Y, Z LOCATION)(TIME) (VALUE)

FIGURE 1

Paper 6

Monitors.

The Storet System is unique as a central repository for data from many sources. It is rightly termed a collection of data bases. The validating tests listed are optional. The value range tests became part of the system in 1975. They consist of user supplied station or run specific value ranges. Application of this logic with a single range to all data accumulated prior to 1975, approximately 30 million data points, identified approximately 45,000 exceptions. To date, these have not been edited.

VALIDATING TESTS IN USE

| | Class | | |
|--|--|--|--|
| | I | II | III |
| <u>DATA BASE</u> | | | |
| <u>USGS 1.</u> | | | |
| Monitors | Syntax, Component Logic, Reference Comparison | Inter-Parameter Relationships | Continuity with time and location |
| Grab Samples | Syntax, Component Logic, Reference Comparison | Balances, Inter-Parameter Relationships | Continuity with location |
| <u>NCC 2.,3.,4.,5</u> <u>Existing</u> | Syntax, Reference Com- parison | Inter-Parameter Relationships | Continuity with time and location |
| Proposed | Syntax, Reference Com- parison | Inter-Parameter Relationships | Continuity with time and location |
| <u>ORSANCO 6.</u> | | | |
| Monitors | Range Checks a. Machine Limits b. Natural Limits | | Range Checks - 95% Confidence Limits |
| Grab Samples | Syntax, Component Logic, Reference Comparison | | Single Range Check |
| <u>STORET 7.</u> | Syntax, Component Logic, Reference Comparison | | Range Check - Station or Run Specific |

HIERARCHAL GLOSSARY

1. Parameter - Name of any measured quantity; i.e., dissolved oxygen, pH, specific conductance, etc.; an independent component of a record.
2. Location - Designation of sampling point in 3 dimensional space, often times an arbitrarily assigned "station number" and a depth. The location consists of 3 independent components of a record.
3. Time - Date and time of sample collection, an independent component of a record.
4. Value - Digital results of a measurement, the dependent component of a record.
5. Record Identity - The set of 5 independent components of a record; parameter, 3 dimensional location and time.
6. Record - A value and associated identity.
7. Sample - Any complete record set with common location and time, parameter is varied. Complete is interpreted as "all which were selected for determination".
8. Validation - Act of testing for correctness.
9. Exception Record - Record which failed some test for validity.
10. Edit - Act of changing a record.
11. Data - Any set of records.
12. Data Base - The set of all historical records.

APPENDIX

Paper 6

REFERENCES

1. United States Geological Survey, "Watstore User's Guide", Reston, VA.
2. Klint, W. E., "Screening Checks Used by the National Climatic Center for Meteorological Data", presented at the Conference on Data Validation, November 4, 1977, Research Triangle Park, NC.
3. Jalickee, J., Sullivan, J., and Rozett, R., "Validation, Compaction, and Analysis of Large Environmental Data Sets", EDS Journal, Environmental Data Service, National Oceanographic and Atmospheric Administration, May 1975, pp. 3-9.
4. Barton, G. and Saxton, D., "The Role of Interactive Computer Systems in Data Processing at CEDDA", EDS, Environmental Data Service, National Oceanographic and Atmospheric Administration, May 1975, pp. 10-14.
5. Good, I. J., "Some Applications of the Singular Decomposition of a Matrix", Technometrics, Vol. II, No. 4, November 1969, pp. 823-831.
6. Ohio River Valley Water Sanitation Commission, Cincinnati, OH.
7. United States Environmental Protection Agency, "Storet Users Manual", Washington, DC.

TEMPERATURE AND DISSOLVED OXYGEN DATA INTERPRETATION

By

Dennis E. Ford*

INTRODUCTION

The data interpretation process is viewed differently by people depending on their particular needs and interests. A statistician may view data interpretation only in terms of statistics while a policy maker may view it as a definite answer to the question: Does a particular water-quality parameter violate standards? In this paper, data interpretation is considered synonymous with a basic understanding of all factors and processes responsible for or influencing the particular phenomenon or parameter of interest. A logical explanation therefore exists for all observed phenomena.

The general theme of this paper will be to identify some of the important factors and processes influencing temperature and dissolved oxygen (DO) phenomena and to illustrate them with field data. This can be a formidable task for water-quality data because of the numerous interactions between the morphometry, physics, chemistry, and biology of a river or reservoir system. All of these interactions must be considered in the interpretation process for all water-quality variables including temperature and DO.

Only reservoirs will be considered in this paper. Since similar processes occur in rivers, some of the information and techniques presented herein will also be applicable for riverine systems. The general order of presentation will be reservoir characteristics, temperature data, DO data, and oxygen demands.

RESERVOIR CHARACTERISTICS

Morphometry - According to recent estimates, there are over 400 authorized CE reservoirs. These reservoirs vary in surface area from approximately 1 km^2 to 10^4 km^2 and in volume from approximately 10^6 m^3 to 10^{11} m^3 . The general shape of most reservoirs is probably best described as dendritic. In such a heterogeneous group, it is unlikely

* Hydrologist, Environmental Laboratory, U. S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi 39180.

that similar processes dominate in all reservoirs, but relationships do appear to exist between basic morphometric parameters such as surface area, volume, depth, drainage basin size, and inflow (e.g., Figures 1-3). These relationships are encouraging because they bring a semblance of order to a diverse group of reservoirs. These morphometric relationships indicate that reservoirs with similar morphology may exhibit other similarities and that relationships with other parameters such as thermocline depth and mixing coefficients may exist (Figures 4 and 5).

An important factor of reservoir morphometry is the depth to length ratio (i.e., H/L). A typical value is 10^{-3} . This ratio implies that problems involving vertical variations may be treated independently from problems involving horizontal variations. It also explains why one-dimensional mathematical reservoir models have been relatively successful in describing reservoir dynamics.

Flow Regime - The flow regime within a reservoir is generally characterized as turbulent and unsteady. Typical Reynolds numbers are always greater than 10^6 , clearly indicating turbulent flow. In stratified reservoirs, where buoyancy forces dominant, the Reynolds number is not a valid criterion for turbulence. Instead, a Richardson number or densimetric Froude number is used. These criteria indicate sporadic turbulence in all but the most stable density gradients. Generally, the turbulence is nonisotropic because of the small H/L. This implies that the reservoir is not homogeneous, and that obtaining representative water-quality samples may be difficult.

The flow regime is unsteady because the controlling or driving forces are unsteady. For example, daily variations in wind, solar radiation, and inflow are shown in Figures 6 and 7. Diurnal, nocturnal, and hourly variations also occur. The response of a lake's water surface temperature and mixed layer* depth to daily variations in solar radiation and wind speed is shown in Figure 6. Under constant forcing, it would take several days for a medium size reservoir to attain steady state conditions. Such steady state conditions are unlikely in reservoirs.

Reservoir ≠ Natural Lake - There are several reasons why a reservoir may differ from a natural lake. Several of these reasons are discussed by Baxter (1977). One obvious reason is some reservoirs have the capability to release water from different depths (i.e., selective or bottom withdrawal) rather than only at the water surface, as in a natural lake. Reservoirs also have shorter residence times, larger drainage basins, larger shoreline development ratios, etc. The impact of each of these differences on the water quality of a reservoir is beyond the scope of this paper; however, at this time it is only necessary to understand that differences do exist and that extrapolating

* The mixed layer is defined later in the text and in Figure 8.

results from natural lakes to reservoirs may be subject to error.

A natural lake may, however, be a simplified subset of a reservoir. In many instances, important processes are much easier to identify in natural lakes since there are fewer complicating factors. Some processes from natural lakes will therefore be used to illustrate important points in this paper.

TEMPERATURE DATA

Of all the water-quality parameters, temperature data is perhaps the easiest to obtain and the most useful. Since all chemical and biological processes are temperature dependent to varying degrees and since these processes are also governed by the degree of stratification and mixing, temperature measurements are essential to any water-quality study. A reservoir is characterized as being of poor or good water quality based on many other variables in addition to temperature.

Thermal Stratification - Before discussing thermal stratification, it is important to review some basic terminology. Several basic limnology books describe stratification in detail (e.g., Wetzel, 1975; Ruttner, 1963; Hutchinson, 1957). The important terms are illustrated in Figure 8. The epilimnion is the upper strata of warm, turbulent water. It is usually characterized by relatively uniform temperatures. The deep, cold, relatively undisturbed region is termed the hypolimnion. Between the epilimnion and hypolimnion is the metalimnion which is characterized by a strong temperature (density) gradient. The plane of inflection or of maximum temperature gradient is termed the thermocline. Other definitions have been proposed for the thermocline (e.g., the $1^{\circ}\text{C}/\text{m}$ criterion), but they are not in wide use and should be avoided. Another term illustrated in Figure 8 is the "mixed layer." The mixed layer is, as implied by the name, the overlying isotropic layer. Since the mixed layer refers to the instantaneous depth of the overlying isotropic layer, it differs from the epilimnion, which is an averaged mixed layer, in two respects. First, its depth is usually less than the depth of the epilimnion. Second, it is much more dynamic than the epilimnion.

There are many other terms used to describe the thermal structure of a lake. These can be found in the above references and are related to the number of turnovers (i.e., periods of complete vertical mixing) occurring within a lake and to the strength of the stratification. For the most part, these terms are of minor importance compared to the difference between holomictic and meromictic lakes. In holomictic lakes, the entire water column completely circulates or turns over. Lakes that cannot circulate completely, and exhibit a deep strata that is perennially stagnant in the water column, are termed meromictic lakes. The reason for this condition may be physical, chemical, or biological. In any case, it can have a profound effect on the

temperature and mixing structure and consequentially on the water quality of a lake. Bottom withdrawal is an effective means to eliminate this undesirable condition.

Factors Affecting Stratification - Since all lakes exhibit some degree of thermal stratification, it is important to investigate the factors responsible for stratification. The principal factors influencing the formation, strength, and extent of stratification are the density of water, solar radiation and the heat budget at the air-water interface, and the mixing resulting from advection and wind-induced phenomena.

It is well known that the density of water varies with temperature (Figure 9). The importance of this variation in determining the distribution of heat within a lake was originally documented by Birge (1910). Two factors are important. First, the maximum density of water occurs at 4°C. Colder or warmer water will be less dense and tend to float on top of the heavier 4°C water. Second and more important, the density of water decreases with an escalating rate with increasing temperatures above 4°C. Therefore, the buoyant forces and resistance to mixing increases with temperature. There is over an order of magnitude difference in the energy requirements to mix a 1°C temperature difference at 25°C than at 5°C.

The energy available to warm the waters of a reservoir ultimately comes from solar radiation which varies seasonally. The seasonal variation of solar radiation follows a sinusoidal curve with a maximum in late June. In addition, diurnal cycles also occur. Water temperatures respond to both of these cycles with a slight delay.

Solar radiation is absorbed at the water surface and selectively with depth depending on the wavelength of the light, properties of the water, and the matter suspended in the water. This absorption is usually assumed to be exponential with depth (i.e., Beer's Law), but surface effects result in minor discrepancies in the top meter or so of a lake (Figure 10). To avoid these discrepancies, it is sometimes assumed, especially by mathematical modelers, that a certain fraction (β) of the solar radiation is absorbed in a surface layer of a specified thickness (e.g., 2 ft (0.6 m)) and the remaining fraction is absorbed exponentially with depth with an extinction coefficient, η . The specific magnitudes of β and η depend on the properties of the water, but there is a definite relationship between β and η (Figure 11). For physical significance, the β and η used in mathematical model studies should reflect this relationship. If field measurements are not available, the extinction coefficient can be estimated directly from Secchi disk measurements (Figure 12). Considering all the uncertainties that go into a Secchi disk measurement, the relationship depicted in Figure 12 is excellent.

Field measurements are also remarkably consistent in predicting the Secchi disk depth to occur at the 10-15 percent light level (Ford, 1976). This fact is useful in calculating extinction coefficients

directly from Secchi disk measurements. For example, assuming exponential absorption with depth, the vertical distribution of light with depth is

$$I_z = I_0 e^{-\eta z}$$

where I_z is the light intensity at depth z

I_0 is the light intensity at the water surface.

Now, if the Secchi disk depth is assumed to occur at the 12.5 percent light level, then I_z/I_0 is 0.125 and the extinction coefficient is given by

$$\eta = 2.1/SD$$

where

SD is the Secchi disk depth in meters

η is the extinction coefficient in 1/m.

The problem with using these estimates of light extinction is the penetration of light is dependent on algae concentrations, suspended solids, etc., and highly variable (Figure 13). Some investigators have been successful in relating Secchi disk depths to a suspended solids concentration (F. Schiebe, personal communication, 1978) and chlorophyll *a* (Carlson, 1977). Comparison of Figures 13 and 14 shows how important light penetration is in determining the depth of the thermocline. The thermocline depth is deeper for the years 1968-72 when the Secchi disk depths are larger.

In contrast to heating, cooling of a water body can only occur at the water surface. It is therefore possible for the surface water to decrease in temperature while deeper water increases in temperature. If the temperature of the surface water drops below the temperature of the deeper water and still remains above 4°C, the water column becomes thermally unstable and natural convection and mixing commences. The convection mixing is termed penetrative convective mixing if it results in the entrainment of water from an underlying region of stable density (temperature) gradient. Otherwise, it is termed nonpenetrative convective mixing and mixing occurs only down to a depth to eliminate density instabilities. It is generally believed that convective mixing in lakes and reservoirs is penetrative although most mathematical models consider it to be nonpenetrative.

Mixing in lakes and reservoirs results from the cumulative effects of inflows, outflows, wind generated currents, surface and internal waves, Langmuir circulation, natural convection, etc. Since the contribution of these phenomena and their interactions to mixing are not completely understood, their cumulative effects are usually lumped into an eddy or effective diffusivity term. This approach is commonly used in mathematical modeling.

The eddy diffusion approach is based on an analogy with molecular diffusion. In molecular diffusion, a property is transferred down a concentration gradient by the random motion of molecules with no overall transport of the fluid. The flux or transport of a property is therefore equal to a molecular diffusion coefficient times a concentration gradient. That is

$$\text{Flux} = K_m \frac{\partial c}{\partial x}$$

where

K_m is the molecular diffusion coefficient

$\frac{\partial c}{\partial x}$ is the concentration gradient.

The eddy diffusion approach also assumes that the turbulent flux of a constituent is proportional to an eddy diffusion coefficient times a gradient. The eddy diffusion coefficients are usually assumed to be much larger than the molecular diffusion coefficient. One important distinction of the molecular diffusion coefficient is a property of the fluid while the eddy diffusion coefficient depends on the flow, density stability, etc., of the fluid. A simple example will illustrate the implications of this fact. It is commonly recognized that the strong gradients in the metalimnion of a lake inhibit turbulence and mixing. Since the eddy diffusion approach assumes the transport of a constituent is proportional to the concentration gradient, mixing or transport will be greatest in the metalimnion where the gradients are the largest. This conclusion is contrary to field observations. Extreme care must therefore be used in analyzing data based on this approach.

Eddy diffusion coefficients can be back calculated from temperature data and the one-dimensional, unsteady, thermal energy equation (i.e., diffusion equation)

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial z} K_z \frac{\partial T}{\partial z} - \frac{\partial H_z}{\partial z}$$

where T is the mean temperature

z is the depth

K_z is the eddy diffusion coefficient

H_z is the internal flux of solar radiation

If this equation is integrated down the water column from depth z to the bottom depth H , then

$$K_z = \left\{ \int_z^H \frac{\partial T}{\partial t} dz - H_z \right\} / \left(-\frac{\partial T}{\partial z} \right).$$

This equation can be solved graphically or with the aid of a computer. Typical results are shown in Figure 15. Variations with time and

directly from Secchi disk measurements. For example, assuming exponential absorption with depth, the vertical distribution of light with depth is

$$I_z = I_0 e^{-\eta z}$$

where I_z is the light intensity at depth z

I_0 is the light intensity at the water surface.

Now, if the Secchi disk depth is assumed to occur at the 12.5 percent light level, then I_z/I_0 is 0.125 and the extinction coefficient is given by

$$\eta = 2.1/SD$$

where

SD is the Secchi disk depth in meters

η is the extinction coefficient in 1/m.

The problem with using these estimates of light extinction is the penetration of light is dependent on algae concentrations, suspended solids, etc., and highly variable (Figure 13). Some investigators have been successful in relating Secchi disk depths to a suspended solids concentration (F. Schiebe, personal communication, 1978) and chlorophyll *a* (Carlson, 1977). Comparison of Figures 13 and 14 shows how important light penetration is in determining the depth of the thermocline. The thermocline depth is deeper for the years 1968-72 when the Secchi disk depths are larger.

In contrast to heating, cooling of a water body can only occur at the water surface. It is therefore possible for the surface water to decrease in temperature while deeper water increases in temperature. If the temperature of the surface water drops below the temperature of the deeper water and still remains above 4°C, the water column becomes thermally unstable and natural convection and mixing commences. The convection mixing is termed penetrative convective mixing if it results in the entrainment of water from an underlying region of stable density (temperature) gradient. Otherwise, it is termed nonpenetrative convective mixing and mixing occurs only down to a depth to eliminate density instabilities. It is generally believed that convective mixing in lakes and reservoirs is penetrative although most mathematical models consider it to be nonpenetrative.

Mixing in lakes and reservoirs results from the cumulative effects of inflows, outflows, wind generated currents, surface and internal waves, Langmuir circulation, natural convection, etc. Since the contribution of these phenomena and their interactions to mixing are not completely understood, their cumulative effects are usually lumped into an eddy or effective diffusivity term. This approach is commonly used in mathematical modeling.

The eddy diffusion approach is based on an analogy with molecular diffusion. In molecular diffusion, a property is transferred down a concentration gradient by the random motion of molecules with no overall transport of the fluid. The flux or transport of a property is therefore equal to a molecular diffusion coefficient times a concentration gradient. That is

$$\text{Flux} = K_m \frac{\partial c}{\partial x}$$

where

K_m is the molecular diffusion coefficient

$\frac{\partial c}{\partial x}$ is the concentration gradient.

The eddy diffusion approach also assumes that the turbulent flux of a constituent is proportional to an eddy diffusion coefficient times a gradient. The eddy diffusion coefficients are usually assumed to be much larger than the molecular diffusion coefficient. One important distinction of the molecular diffusion coefficient is a property of the fluid while the eddy diffusion coefficient depends on the flow, density stability, etc., of the fluid. A simple example will illustrate the implications of this fact. It is commonly recognized that the strong gradients in the metalimnion of a lake inhibit turbulence and mixing. Since the eddy diffusion approach assumes the transport of a constituent is proportional to the concentration gradient, mixing or transport will be greatest in the metalimnion where the gradients are the largest. This conclusion is contrary to field observations. Extreme care must therefore be used in analyzing data based on this approach.

Eddy diffusion coefficients can be back calculated from temperature data and the one-dimensional, unsteady, thermal energy equation (i.e., diffusion equation)

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial z} K_z \frac{\partial T}{\partial z} - \frac{\partial H_z}{\partial z}$$

where T is the mean temperature

z is the depth

K_z is the eddy diffusion coefficient

H_z is the internal flux of solar radiation

If this equation is integrated down the water column from depth z to the bottom depth H , then

$$K_z = \left\{ \int_z^H \frac{\partial T}{\partial t} dz - H_z \right\} / \left(-\frac{\partial T}{\partial z} \right).$$

This equation can be solved graphically or with the aid of a computer. Typical results are shown in Figure 15. Variations with time and

depth (i.e., stratification) are evident. If daily temperature data are used, then the calculated eddy diffusion coefficients can vary by several orders of magnitude from one day to the next. It should be noted that this approach does not consider horizontal advection. Since some reservoirs are dominated by horizontal advection, the use of this approach may result in unrealistic estimates for eddy diffusion coefficients.

One application of eddy diffusion coefficients is to calculate the flux of a nutrient (e.g., phosphorus) from the metalimnion to the epilimnion. The transport of mass and momentum are not necessarily equal to the transport of thermal energy. Since most diffusivities are based on the transport of thermal energy (i.e., back calculated from temperature data), their application to other constituents such as dissolved oxygen or nutrients may be questionable.

Recent sensitivity analyses on an ecological model at the Waterways Experiment Station (WES) (Thornton and Lessem, 1976) showed the diffusion coefficient to be a highly sensitive variable. Perturbations of the diffusion coefficient that had minimal effect on the temperature structure had dramatic effects on the algae concentrations. One reason for this sensitivity is differences between concentration gradients. Although the physical location (depth) of maximum gradients may be similar for different constituents (e.g., the metalimnion), the magnitudes of the gradients are not similar. For example, in the metalimnion of a lake, temperatures may vary between 10° and 20°C while nutrient concentrations may vary over two orders of magnitude. The use of a single constant eddy diffusion coefficient to determine the transport of temperature and nutrient in this example would result in an excessive transport of the nutrient.

Seasonal Temperature Cycle - To see how the above factors interact to form a seasonal temperature cycle, a natural lake will be used as a simple example (Figure 16) to minimize the effects of inflow and outflow. In Figure 16, the ice cover on the lake went out on 17 April. Thereafter, the entire water column mixed for several days (i.e., period of spring turnover). The density differences at the low temperature were not sufficient to prevent complete mixing. As the water column warmed, density differences increased and it became more difficult to mix the entire water column. Stratification started to form at the bottom of the lake because the density differences and resulting buoyancy forces were small compared to the kinetic energy input (i.e., wind). As the solar radiation increased; water temperatures increased; density differences increased; and the thermocline moved upward because the kinetic energy input could not overcome the ever increasing buoyancy forces. The minimum thermocline depth was usually achieved sometime around the time of summer solstice or time of maximum heat input. Once stratification formed, the hypolimnion temperature remained relatively constant until fall overturn. Fall overturn occurred when the surface temperature approached the hypolimnion temperature.

Generally, the slope of the isotherms indicates the degree of mixing; however, this assumes no internal heating from solar radiation occurs. The relatively flat isotherms in the metalimnion during mid-summer indicated very little mixing in this region (Figure 16). Similarly, the steep slope of the 14° and 16°C isotherms during late May and early June should indicate intensive mixing. During this period, however, the Secchi disk depth increased from 2 m to 6 m, indicating the heating was due to internal absorption of solar radiation and not mixing (Figure 17).

The increased mixing resulting from inflows and outflows in reservoirs can result in deviations from the above example. For instance, periods of overturn may be extended and become more frequent, and the slope of the isotherms may be increased. If bottom withdrawal is used, hypolimnetic temperatures may increase. For example, in Figure 18, the hypolimnetic temperatures in Beltzville Lake in 1976 remained relatively constant when only small amounts of water were released through the lower ports. In contrast, in 1972, when large quantities of water were released through the flood gates in response to Hurricane Agnes, the hypolimnetic temperatures increased. Man-induced perturbations such as power operations can also increase mixing and hypolimnion temperatures (Figure 19).

Factors Affecting Data Interpretation - Several other factors need to be considered in temperature data interpretation. First, mixing and entrainment are not gradual continuous processes as indicated by the seasonal temperature structure in Figure 16 and by the eddy diffusion approach. They are actually dynamic processes resulting from the interactions and variations of solar radiation, strong winds, inflows, power operations, etc. Examples showing the extent of mixing resulting from several hours of strong winds are given in Figure 20. Entrainment of nutrients from the metalimnion into the euphotic zone can therefore be considered as an instantaneous process rather than a gradual continuous process.

In addition to the seasonal cycle in temperature, there are also diel and synoptic cycles in temperature (Figures 21 and 22). These result from variations in solar radiation and mixing. These temporal variations can be important in the verification of mathematical models and in data interpretation. Since most mathematical models use a daily time step, predictions represent a full day of heating and cooling. In contrast, field measurements are usually taken during daylight hours and frequently in mid-afternoon when diurnal heating is maximum. Diel variations can be as large as 7°C or more, but are typically the order of 1° or 2°C. The magnitude of the variation will depend on the size and depth of the lake, among other factors.

Horizontal variations also occur. These can result from differential heating, inflow, or mixing. Differential heating occurs when the smaller volume of water in the littoral zones and headwaters of

impoundment warms or cools more rapidly than the open water regions. In large lakes, this phenomenon is significant and results in the formation of thermal bars. Similarly, rivers flowing into a reservoir may be of different temperatures creating longitudinal variations. Examples of horizontal variations are shown in Figure 23. Horizontal variations typically create temperature differences of 1° or 2°C or more.

Consideration must be given to the above mentioned variations and to the physical processes responsible for them in data interpretation. If discrepancies in data are not explainable, then the data could be erroneous. An example of erroneous data is shown in Figures 18 and 24 for Cave Run Lake. The hypolimnetic temperatures for Julian day 213 are unrealistic with respect to days 184 and 219. Examination of the meteorological data for this period confirmed that there were no extreme events that could have possibly caused such a temperature deviation.

DISSOLVED OXYGEN DATA

Two basic reasons for measuring dissolved oxygen in lakes and reservoirs are aesthetics and aquatic life. DO is required in reservoirs to prevent the onset of septic conditions and the accompanying malodorous emissions. The development of carbon dioxide, hydrogen sulfide, and methane in the sediments under anaerobic conditions can loosen sludge which can float to the water surface and form an unsightly scum. DO is also essential to the metabolism of aquatic life. EPA (1976) recommends a minimum of 5 mg/l DO to maintain good fish populations. DO is an important water-quality parameter that influences many other chemical and biological processes in the reservoir.

The sources of DO to a reservoir are photosynthesis, atmospheric reaeration, and hydromechanical. DO sinks are community respiration and oxidation of organic matter. The magnitude of these sources and sinks will depend on reservoir morphology, water temperature, meteorology, the amount and characteristics of the oxidizable substance, etc. For lakes and reservoirs, the most important sources and sinks are probably photosynthesis, community respiration, and autochthonous organic matter (see Table I and Ruttner, 1963).

Before discussing DO distributions in reservoirs, the solubility properties of DO in water need to be reviewed. DO is poorly soluble in water, because it does not react chemically with water and because air contains only about 21 percent oxygen by volume. The solubility of DO is proportional to its partial pressure. It therefore, varies with pressure and temperature (Figure 25). Potential problems with low DO can occur at low pressure (i.e., high altitudes) and at high temperatures. Other parameters such as salinity can reduce the solubility of DO.

TABLE I

Average DO Budget in Lake Keystone on CLOUDLESS DAYS
(after Eley, 1970)

\bar{X} Gains on Clear Days (metric tons O_2)

| | |
|----------------------------------|--|
| Production | + 976.1 day ⁻¹ |
| Diffusion in (\bar{X} k=1.49) | + 42.0 day ⁻¹ |
| Cimarron R. inflow | + 6.8 day ⁻¹ |
| | <hr/> |
| | + 1024.9 +1024.9 day ⁻¹ |

\bar{X} Losses on Clear Days (metric tons O_2)

| | |
|--|--|
| Respiration | - 787.8 day ⁻¹ |
| Diffusion out (\bar{X} k = 1.49) | - 194.8 day ⁻¹ |
| Reservoir discharge | - 1.4 day ⁻¹ |
| | <hr/> |
| | - 984.0 - <u>984.0</u> day ⁻¹ |
| \bar{X} Calculated Net Gain + 40.9 day ⁻¹ | |

DO Distributions - Oligotrophic lakes are typically characterized by low nutrient concentrations and low production of organic matter. The DO distribution is controlled by physical processes such as temperature, pressure, and stratification instead of biological processes. The resulting DO profile is called an orthograde profile. In an idealized orthograde profile, DO is 100 percent saturated and its distribution will depend only on pressure and temperature (Figure 26). Profiles of this type are very seldom found in nature. Usually the DO concentrations in an oligotrophic lake are constant throughout the water column.

In contrast to oligotrophic lakes, eutrophic lakes are characterized by high nutrient concentrations and high organic production. The DO in the hypolimnion is depleted by the oxidation of the settling organic matter and by oxidation at the sediment-water interface. The density gradients in the metalimnion prevent renewal of DO by mixing and light is usually not sufficient for renewal by photosynthesis. The resulting DO profile is called clinograde and it generally follows the shape of the temperature profile (Figure 27). All things being equal, insight into the productivity of a lake can be obtained by examining a DO profile to determine if it is orthograde or clinograde.

Power operations can also affect DO distributions. DO might be expected to improve with the increased mixing from power operations, but this is not necessarily true. In Figure 19, power operations did increase the temperatures in the hypolimnion, but DO did not improve. This example reaffirms the point that estimates of mixing from temperature profiles alone are not sufficient to answer all water-quality questions.

DO Profile Variations - Not all DO profiles are purely orthograde or clinograde. Some exhibit maxima or minima in the metalimnion. The maxima are due to photosynthesis by algae and aquatic plants. The high stability of the metalimnion prevents the DO from diffusing. DO profiles containing a maxima are called positive heterograde curves.

Negative heterograde curves denote metalimnetic DO minima. The minima can result from the accumulation of oxidizable material, basin morphology, zooplankton, and the temperature gradient alone. The sources of the oxidizable material accumulating in the metalimnion can be settling from above (autochthonous) or density currents (allochthonous). In reservoirs, both are possible. The exact source can only be determined by field studies. Attempts at such studies are described by Gordon and Skelton (1977) and Drury and Gearheart (1975). It is possible that allochthonous demands will dominate in the headwaters and autochthonous demands will dominate near the dam. The basin morphology is important if there are large flats of sediment in the metalimnion. Oxidation at the sediment-water interface will create the minima.

Temperature is important because the solubility of DO varies with temperature. Ruttner (1963) argued that the temperature gradient

alone can account for the metalimnetic oxygen minima because oxidation will precede much faster in the epilimnion and metalimnion than in the hypolimnion due to the warmer temperatures. In addition, DO will increase in the epilimnion from transfer through the air-water interface, but not in the metalimnion and hypolimnion because mixing is inhibited. The net result is a metalimnetic minima.

Wetzel (1975) stated metalimnetic maxima are much more prevalent in lakes than metalimnetic minima. Yet, reservoir data (e.g., Figures 19 and 26), continually show metalimnetic minima. There are two possible reasons for this difference. First, in a reservoir, sufficient quantities of allochthonous material may be entering the metalimnion via density currents causing the minima. Second, in natural lakes there appears to be a relationship between the depth of the thermocline and light penetration. Photosynthesis can still occur in the metalimnia causing DO maxima. However, in a reservoir this may not be the case. The increased turbidity from inflows and the increased mixing from advection will result in a deeper epilimnion and may keep light from penetrating into the metalimnion. The result would be a DO minima instead of a maxima.

Seasonal Variations - The seasonal variation in DO as measured in DeGray Lake for three consecutive years is shown in Figure 28. During periods of turnover, DO was distributed uniformly throughout the water column. Both metalimnetic minima and maxima occurred. The minima were prevalent during the summer for the three years while the maxima were not so prevalent and occurred only during the spring. The general increase in DO from 1974 to 1976 illustrates the improvement in water quality that can be expected as a new impoundment progresses through its initial transitory period.

Temporal and Horizontal DO Variations - DO, as temperature, exhibits diel and synoptic variations (Figures 21 and 29). Both examples are related to solar radiation (i.e., clear weather) and the resultant biological activity (photosynthesis and respiration). If temperature was the dominant factor, then the diel DO variation in Figure 21 would be inversely proportional to temperature as implied in Figure 25.

According to Wetzel (1975), the oxygen regime in the littoral zone is totally different from the pelagic (open water) zone. Two reasons for this difference are the large number of aquatic plants and the smaller volume of water found in the littoral zone. For these two reasons, the littoral zone is usually characterized by a more pronounced diel cycle, which in turn, creates horizontal DO variations between the littoral and pelagic regions of a lake. Horizontal DO variations are also created at the end of the growing season when large populations of aquatic plants die and decompose. Since reservoirs characteristically have large shoreline development ratios (i.e., extensive littoral zones), the oxygen regime in the littoral zone can play a significant role in the total oxygen budget of a reservoir.

Allochthonous inflows to a reservoir can also act as a source or sink of DO and thereby create horizontal variations. The oxygen demand will be greater in the headwater regions resulting in lower DO.

OXYGEN DEMANDS

Three general classes of matter that exhibit oxygen demands in reservoirs are carbonaceous material, oxidizable nitrogen, and chemical reducing compounds. The impact of each of these on the DO distribution will depend on many things including the amount of oxidizable material, the settling rate of the oxidizable material, the water temperature, and the bottom profile and depth of the lake. Methods to measure these oxygen demands in reservoirs include biochemical oxygen demand (BOD), chemical oxygen demand (COD), and total organic carbon (TOC), among others.

BOD - BOD is defined as the amount of DO required by bacteria to stabilize decomposable organic matter under aerobic conditions. It has been widely used to determine the pollution strength of domestic and industrial wastes.

The BOD test is a procedure that requires dilute wastes, no toxic substances, accessory nutrients, and a seed consisting of a group of diverse organisms. The reaction is assumed to be complete in 20 days although it theoretically goes to infinity. It is usually measured, however, in 5 days when only a portion of the total BOD is exerted. The precise portion depends on the character of the seed and the nature of the organic matter but it is typically 70 to 80 percent of the total BOD. The total or ultimate BOD is estimated from the 5-day BOD assuming a first order reaction with a constant rate coefficient. For more specific information on the BOD test, see Standard Methods and Sawyer and McCarty (1967).

An idealized BOD reaction is shown in Figure 30. The initial demand is due to carbonaceous material. After about 8 days significant nitrification begins. The 5-day BOD was selected to avoid the oxygen demand from nitrification. The nitrification demand can be measured by measuring the various forms of nitrogen and using stoichiometry to calculate the demand.

Since water samples from rivers, lakes, and reservoirs contain significant populations of nitrifying organisms, nitrification may begin early and thereby affect the interpretation of the 5-day BOD. Although the action of nitrifying bacteria can be arrested by inhibiting agents (e.g., methylene blue) or by reducing populations through pasteurization, chlorination, or acid treatment, there is not a **standard** procedure for measuring the carbonaceous BOD in natural waters (Sawyer and McCarty, 1967). In addition, any type of algae growth in such waters makes BOD data difficult to interpret.

In interpreting BOD data, the assumptions of the BOD test need to be considered. The first assumption is the reaction is first order. Both nitrification and protozoa that eat bacteria can cause deviations from a first order reaction. Second, it is assumed that the rate constant is truly a constant. Table II shows that this is not always true. The constant will depend on the nature of the organic material and the ability of the organisms to utilize organic matter. Since these two assumptions are not always met, extrapolating 5-day BOD to total BOD may be subject to error.

COD - The COD is the oxygen required for oxidation of organic matter to carbon dioxide (CO_2) and water. It is based on the fact that most organic compounds can be oxidized by the action of strong oxidizing agents under acid conditions. A complete description of the COD test is found in Sawyer and McCarty (1967) and Standard Methods.

The disadvantages of the COD are:

(1) Organic matter is converted to CO_2 and water regardless of the biological assimilability of the substance. The COD is usually greater than the BOD (Table II).

(2) The procedure can reduce some inorganic ions causing erroneously high results.

(3) There is no evidence of the rate at which biological active material will stabilize.

The advantages of the COD are that it requires a short time for analysis and it is not subject to toxic conditions.

TOC - The amount of carbonaceous material in a water sample can be obtained from a total carbon analyzer (i.e., infrared spectrophotometer). Through appropriate procedures (see Standard Methods), the organic carbon can be separated from the inorganic carbon to obtain TOC. It is also possible to separate TOC into dissolved organic carbon (DOC) and particulate organic carbon (POC) by filtration, settling, or centrifugation.

The advantage of TOC measurements is that only carbonaceous material is considered. It is also possible to relate TOC and POC to plankton and zooplankton (Figure 31).

The disadvantages of using TOC measurements for estimating oxygen demand are that not all of the carbon may be biologically oxidizable and that there is no evidence on the rate of oxidation.

Discussion - Typical BOD and COD temporal and spatial variations are given in Table II for Lake Keystone. The values decrease from the headwaters (Station I) to the dam (Station IV) indicating the decay of allochthonous material. In addition, the COD is always

TABLE II
SPATIAL AND TEMPORAL VARIATION OF MEAN CONCENTRATIONS OF
BOD AND COD AND MEAN BOD REACTION CONSTANTS
IN LAKE KEYSTONE DURING 1966-67 (after Eley, 1970)

| | 5-day BOD | 20-day BOD | BOD k | COD |
|-------------|--------------|---------------|----------|-----|
| Station I | 3.8 | 11.2 | 0.035 | 176 |
| Station II | 2.5 | 6.6 | 0.041 | 46 |
| Station III | 1.4 | 5.2 | 0.029 | 28 |
| Station IV | 1.1 | 4.5 | 0.026 | 18 |
| 8/1/66 | 1.4 | | | |
| 9/24/66 | 2.1 | 4.7 | | 33 |
| 11/19/66 | 3.1 | 7.9 | 0.042 | 42 |
| 12/17/66 | 2.8 | 8.0 | 0.038 | 48 |
| 1/28/67 | 3.6 | 11.6 | 0.031 | 163 |
| 3/29/67 | 2.8 | 9.2 | 0.028 | 103 |
| 6/4/67 | 1.6 | 7.5 | 0.019 | 36 |
| 7/23/67 | 1.2 | 4.7 | 0.026 | 22 |
| 8/24/67 | 1.1 | 3.1 | 0.042 | 29 |

greater than the BOD values because it is a measure of all oxidizable material, not just the biologically oxidizable material.

In the section on DO, it was stated that the DO budget in lakes and reservoirs was dominated by photosynthesis and community respiration. If this statement is correct, then these processes must be considered when evaluating the oxygen demand in reservoirs. Since the BOD, COD, and TOC do not include the effects of photosynthesis and community respiration, their value for estimating the oxygen demands in reservoirs is questionable. It should also be remembered that the BOD and COD tests are primarily intended to measure the pollution strength of wastes, not the oxygen demands in the reservoirs.

There is not single method available to estimate the oxygen demand in a reservoir. A combination of BOD tests and TOC measurements provides useful information. The BOD will give a measure of the biologically oxidizable organic matter, provided nitrification is arrested. TOC measurements estimate the total quantity of dissolved and particulate organic matter. Together they can be used to estimate biologically resistant organic matter. COD measurements are totally worthless in studying reservoirs because they contribute nothing to the understanding of oxygen dynamics.

Perhaps the best procedure for evaluating oxygen budgets and demands in reservoirs is a time series of DO measurements. All naturally occurring phenomena are therefore taken into consideration in the proper proportions. Measurements of this type are sometimes referred to as oxygen deficits.

SUMMARY

The subject of temperature and DO data interpretation was discussed from the point of view that all data are explainable. This approach requires a complete understanding of the numerous interactions between the morphometry, physics, chemistry, and biology of the system. It is only with this complete understanding that a reservoir can be managed effectively with respect to water quality.

ACKNOWLEDGEMENTS

The review comments of Dr. Kent Thornton, Mr. Joe Carroll, and Mr. Ross Hall, Ecosystem Research and Simulation Division (ERSD), Environmental Laboratory, are greatly appreciated. The ERSD secretarial staff is thanked for typing this manuscript.

REFERENCES

- American Public Health Association. 1971. Standard Methods for the Examination of Water and Wastewater (13th Edition). American Public Health Association, New York.
- Baxter, R. M. 1977. Environmental effects of dams and impoundments. *Ann. Rev. Ecol. Syst.* 8:255-283.
- Birge, E. A. 1910. An unregarded factor in the lake temperatures. *Trans. Wis. Acad. Sci.* 16:989-1005.
- Carlson, R. E. 1977. A trophic state index for lakes. *Limnol. Oceanogr.* 22(2):361-369.
- Culver, D. A. 1975. Physical, chemical, and biological factors in the initiation and destruction of biogenic meromixis in a soft water lake. *Verh. Internat. Verein. Limnol.* 19:776-783.
- Drury, D. D. and R. A. Gearheart. 1975. Bacterial-population dynamics and dissolved oxygen minimum. *Jour. AWWA.* 67(3):154-158.
- Eley, R. L. 1970. Physiochemical limnology and community metabolism of Keystone Reservoir, Oklahoma. PhD Thesis. Oklahoma State University, Stillwater, Oklahoma.
- Environmental Protection Agency. 1976. Quality Criteria for Water, Washington, D.C.
- Faust, A. R. 1972. Continuous monitoring of dissolved oxygen concentration and temperature at multiple depths in a reservoir. MS Thesis. Oklahoma State University, Stillwater, Oklahoma.
- Ford, D. E. 1976. Water temperature dynamics of dimictic lakes: Analysis and predictions using integral energy concepts. PhD Thesis. University of Minnesota, Minneapolis, Minnesota.
- Gordon, J. A. and B. A. Skelton. 1977. Reservoir metalimnion oxygen demands. *Jour. Envir. Eng. Div. ASCE.* 103:1001-1011
- Hutchinson, G. E. 1957. A Treatise on Limnology, Vol. 1. John Wiley and Sons, Inc., New York.
- Ruttner, Franz. 1963. Fundamentals of Limnology (3rd Edition). University of Toronto Press, Toronto.
- Sawyer, C. N. and P. L. McCarty. 1967. Chemistry for Sanitary Engineers. (2nd Edition). McGraw-Hill Book Co., New York.

Thornton, K. W. and A. S. Lessem. 1976. Sensitivity analysis of the
Water Quality for River-Reservoir Systems Model. WES MP Y-76-4.

Wetzel, R. G. 1975. Limnology. W. B. Saunders Company, Philadelphia.

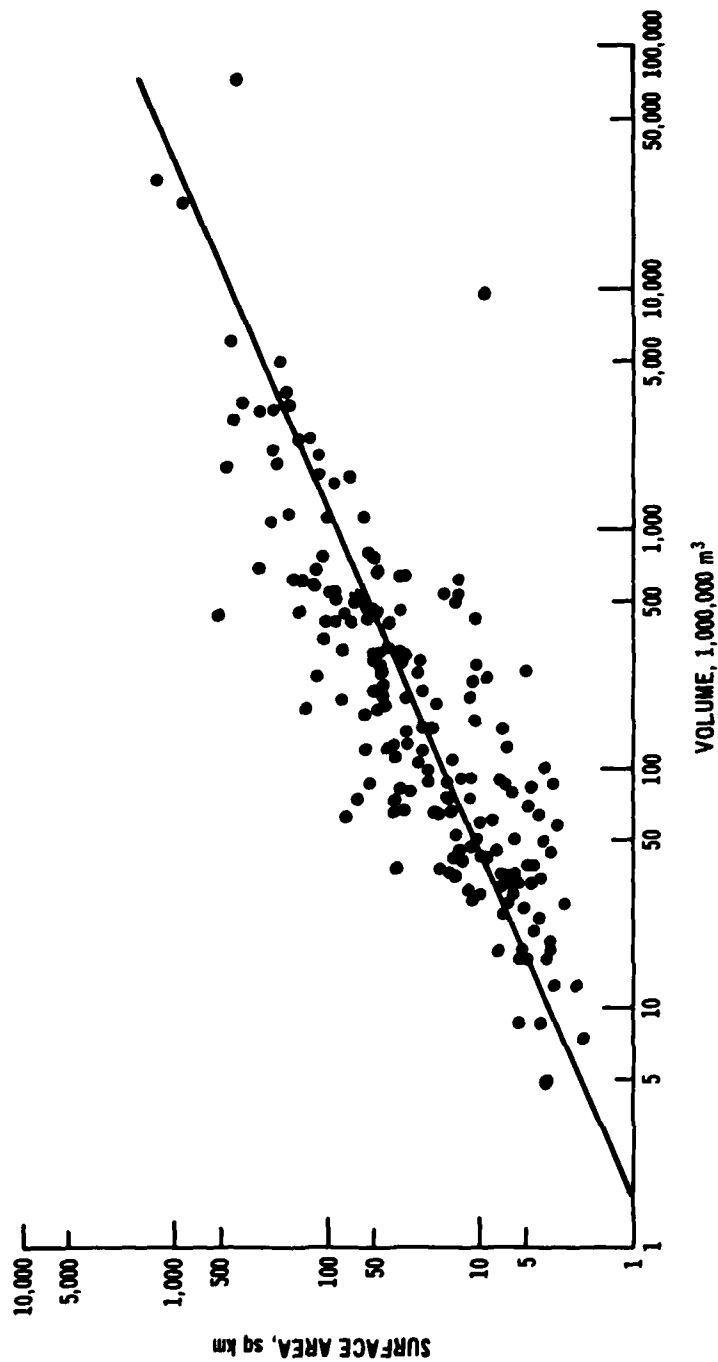


Figure 1. Relationship between surface area and volume for 187 CE impoundments.

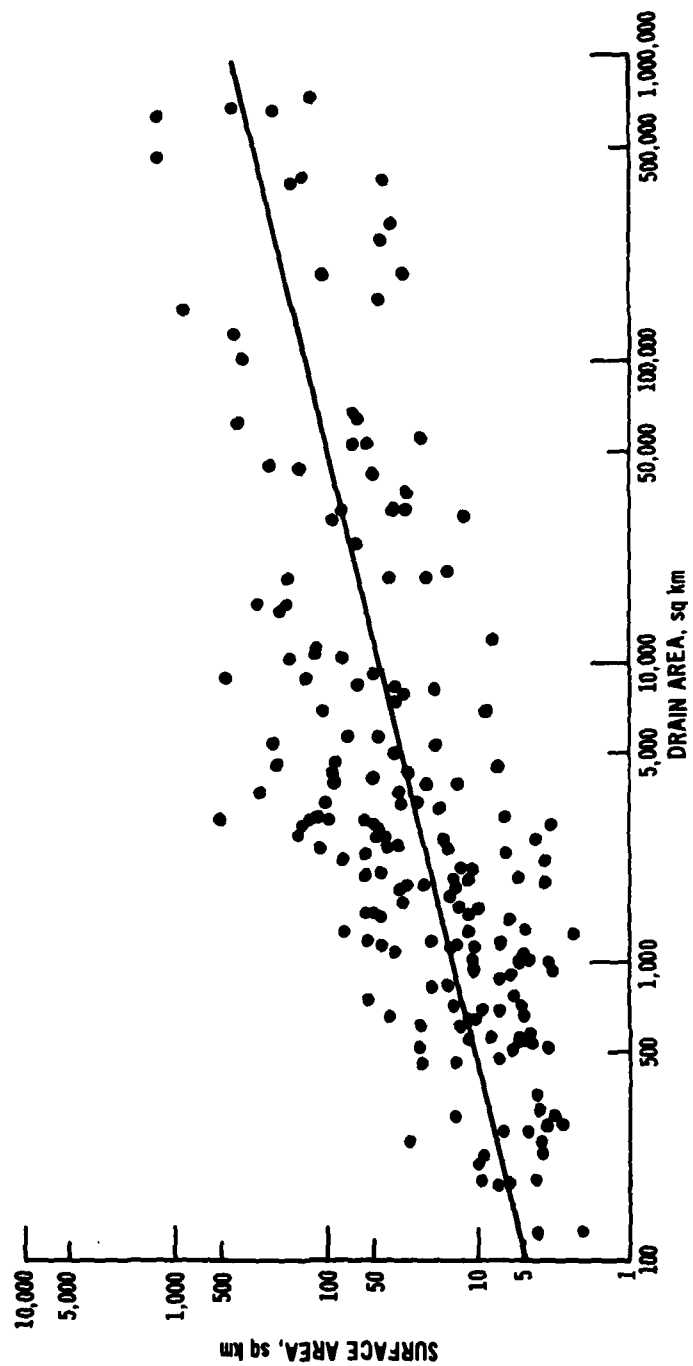


Figure 2. Relationship between surface area and drainage area for 187 CE impoundments.

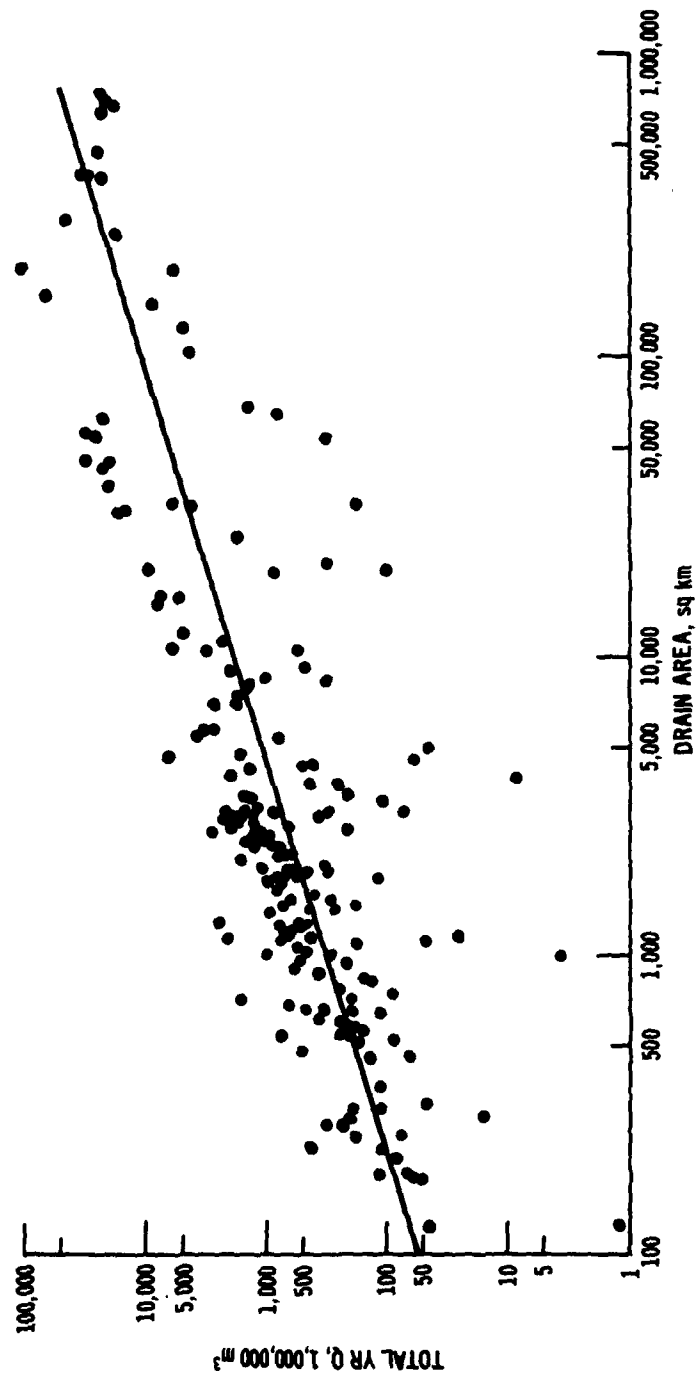


Figure 3. Relationship between total annual flow and drainage area for 187 CE impoundments.

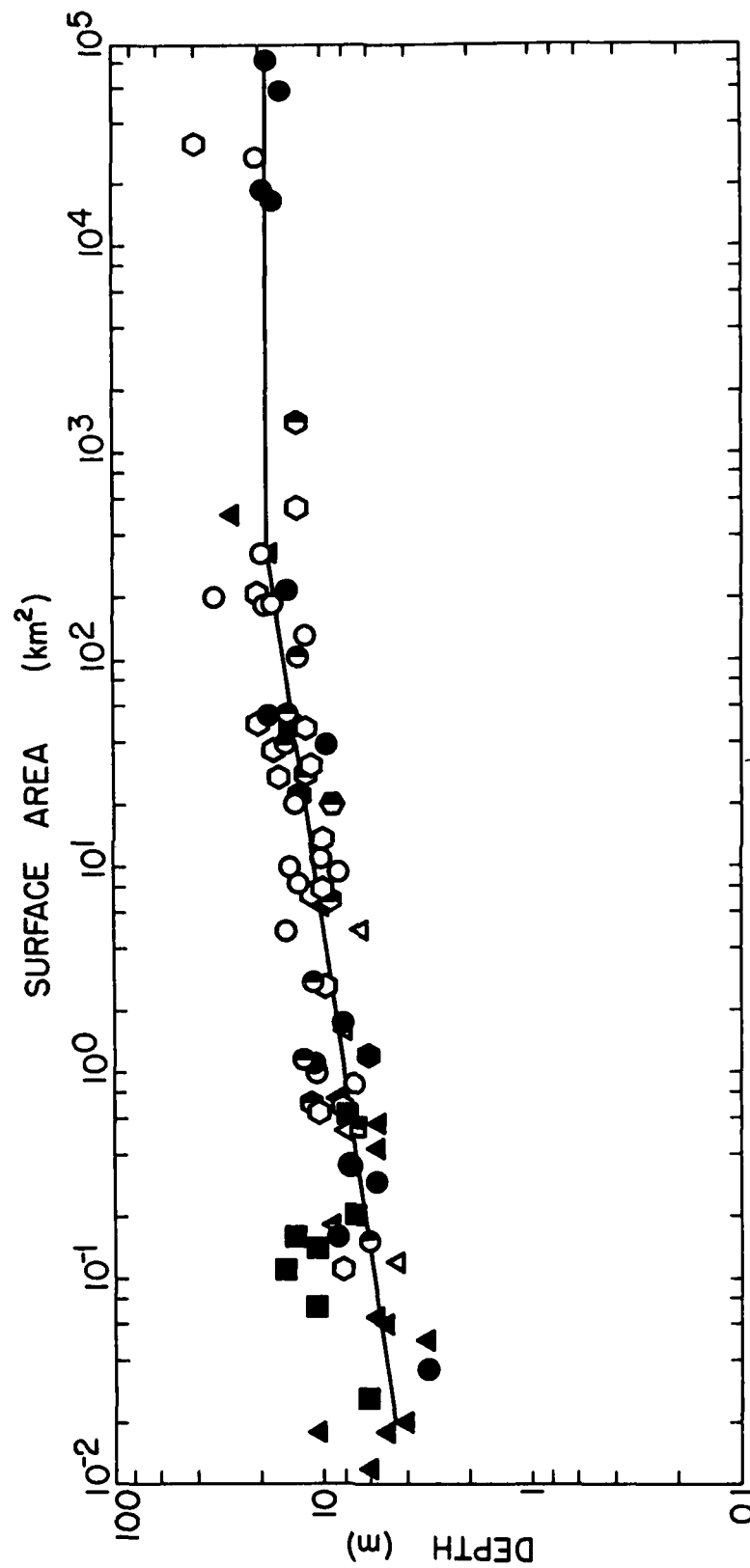


Figure 4. Thermocline depth as a function of surface area. (after Ford, 1976)

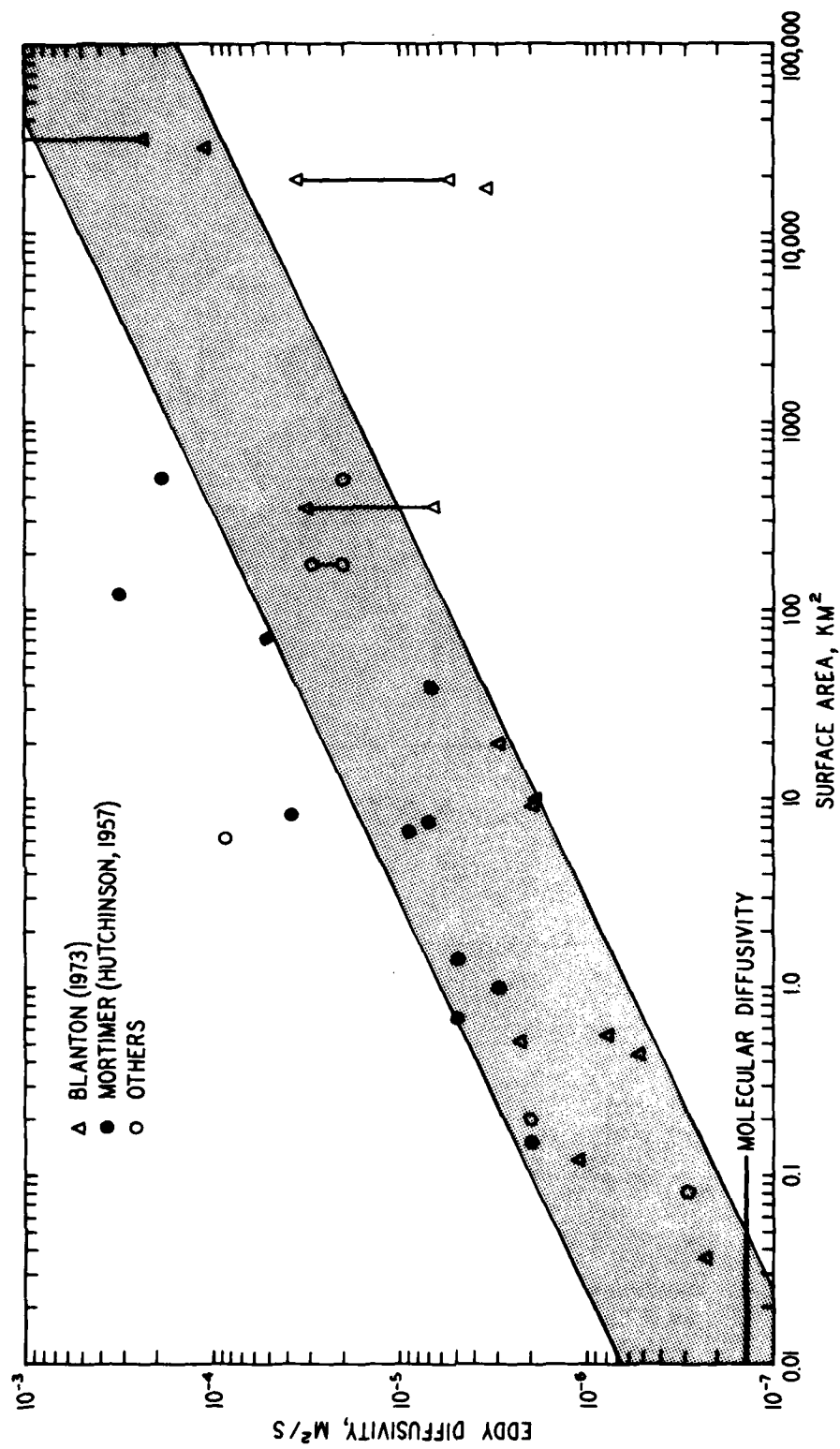


Figure 5. Increase in hypolimnetic eddy diffusivities with lake surface area.

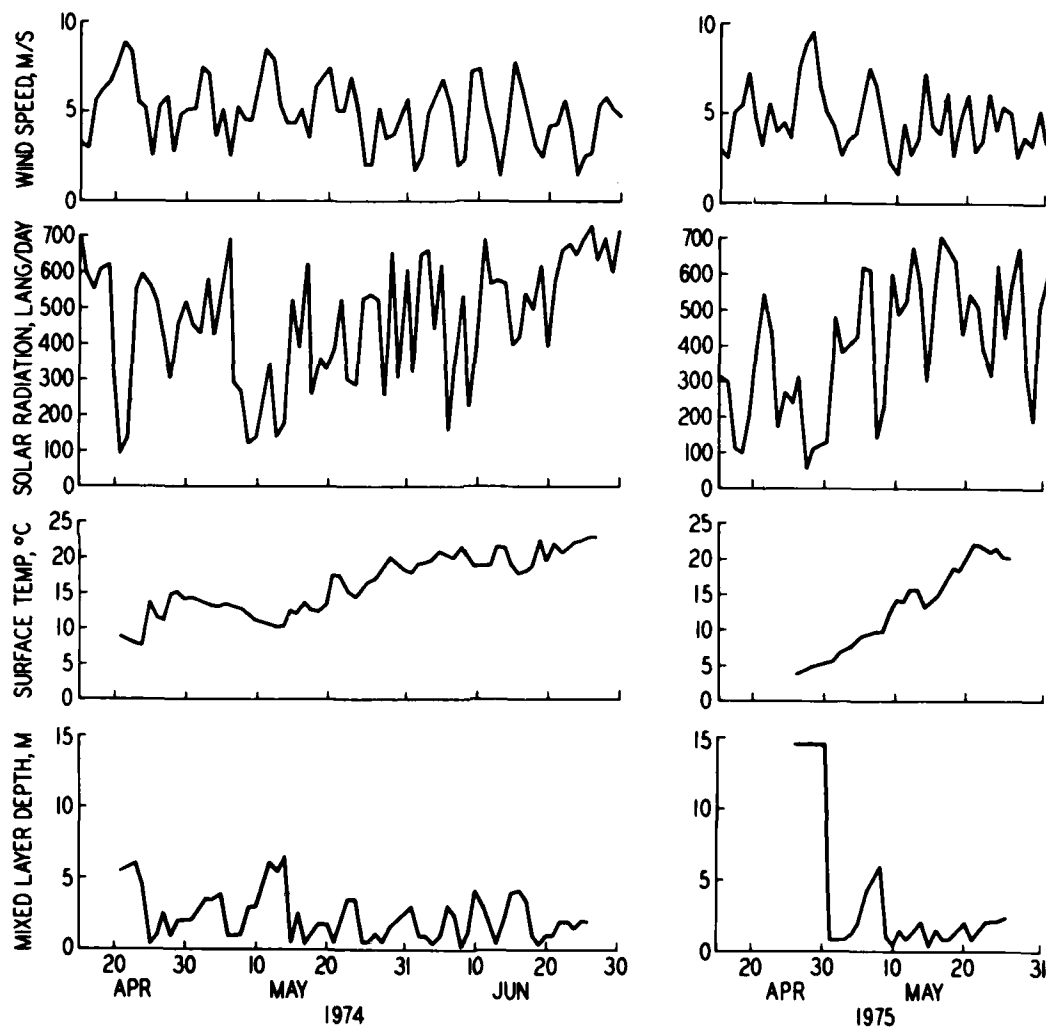
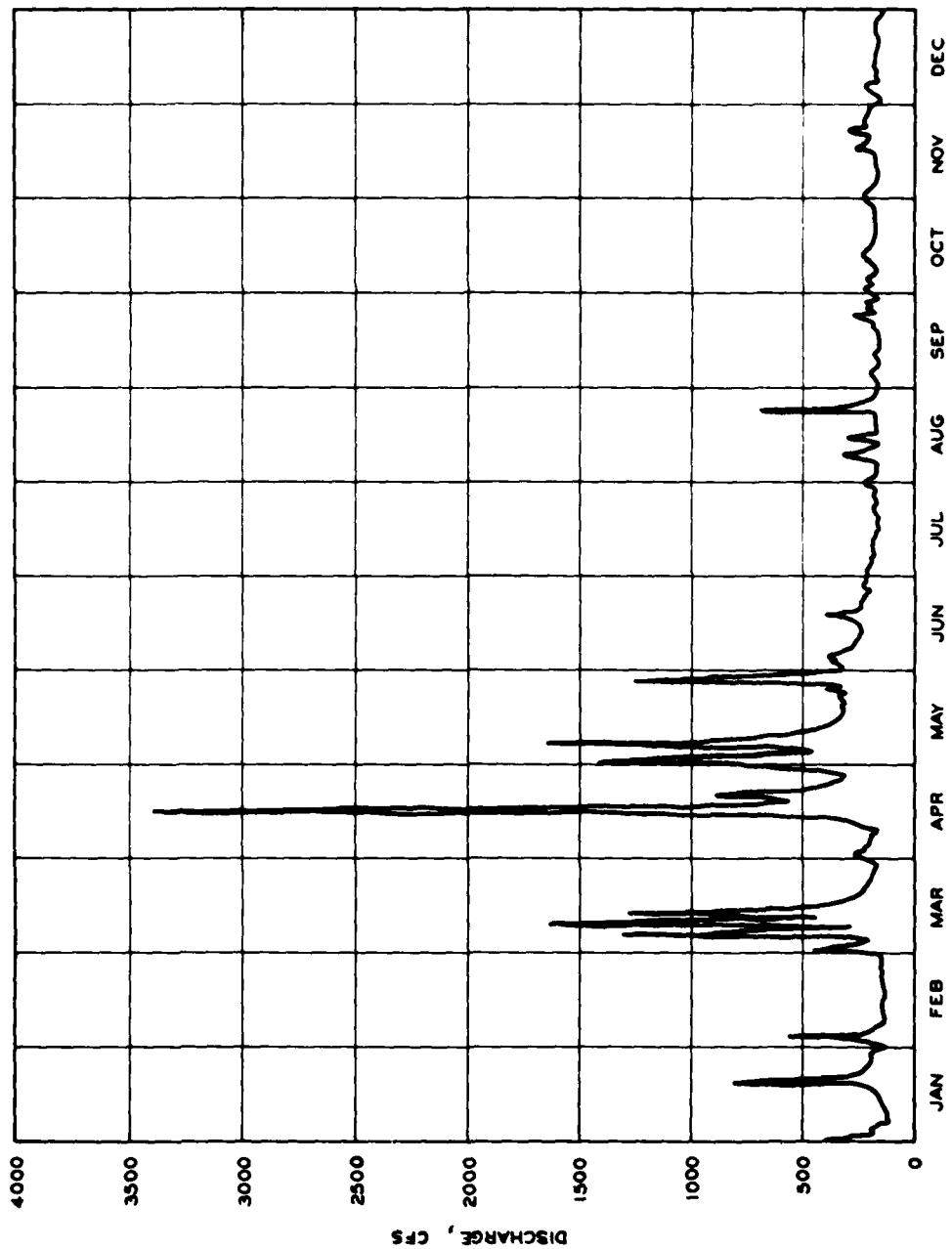


Figure 6. Response of surface temperatures and mixed layer depths in McCarrons Lake, Minnesota, to daily changes in solar radiation and wind speed.



KICKAPOO RIVER DISCHARGE AT LAFARGE, WIS FOR 1973

Figure 7. Typical variations in river flow

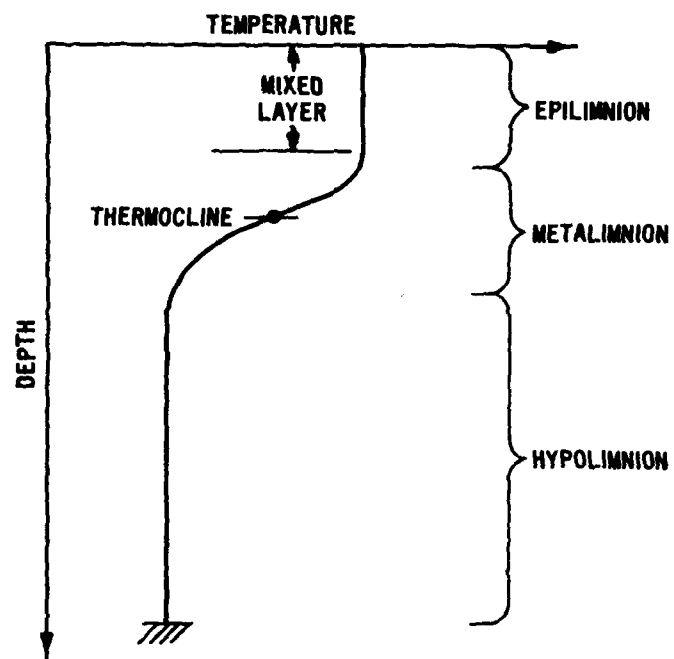


Figure 8. Basic terminology for thermal stratification.

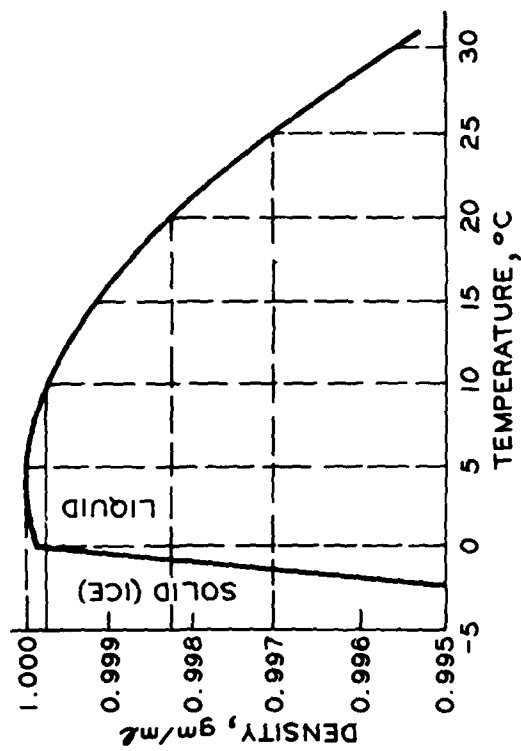


Figure 9. Variation in the density of water with temperatures.

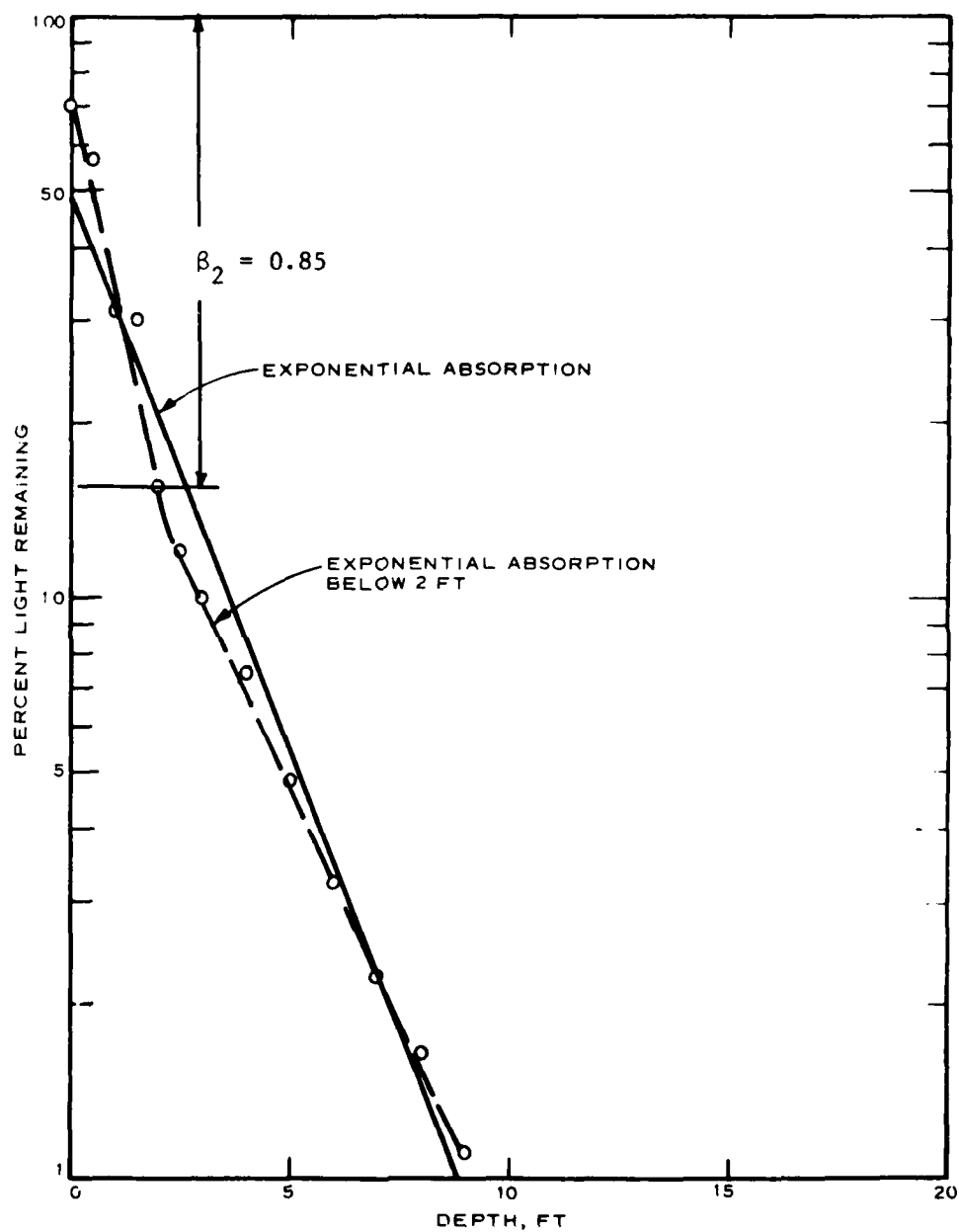


Figure 10. Deviation from exponential absorption in the surface layer, C. J. Brown Lake, Ohio.

AD-A101 802

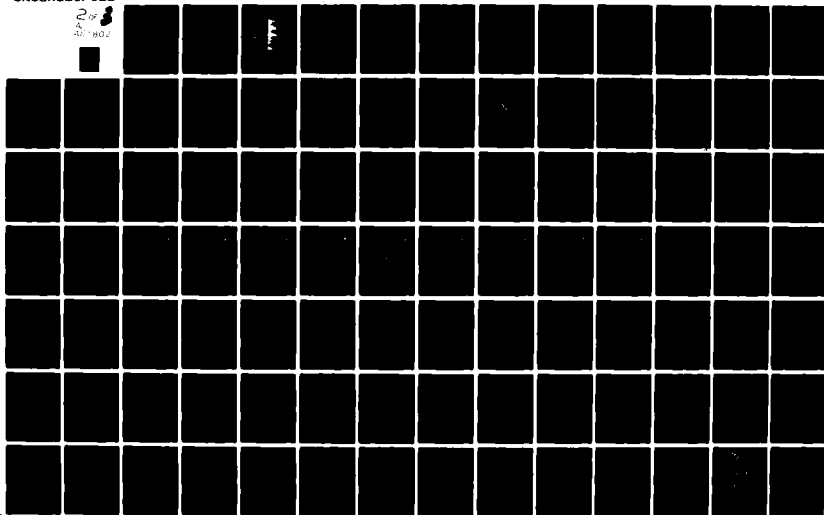
CORPS OF ENGINEERS WASHINGTON DC
PROCEEDINGS OF A SEMINAR ON WATER QUALITY DATA INTERPRETATION, ETC(U)
1978

F/G 13/2

UNCLASSIFIED

NL

2 of 3
AD-A101 802



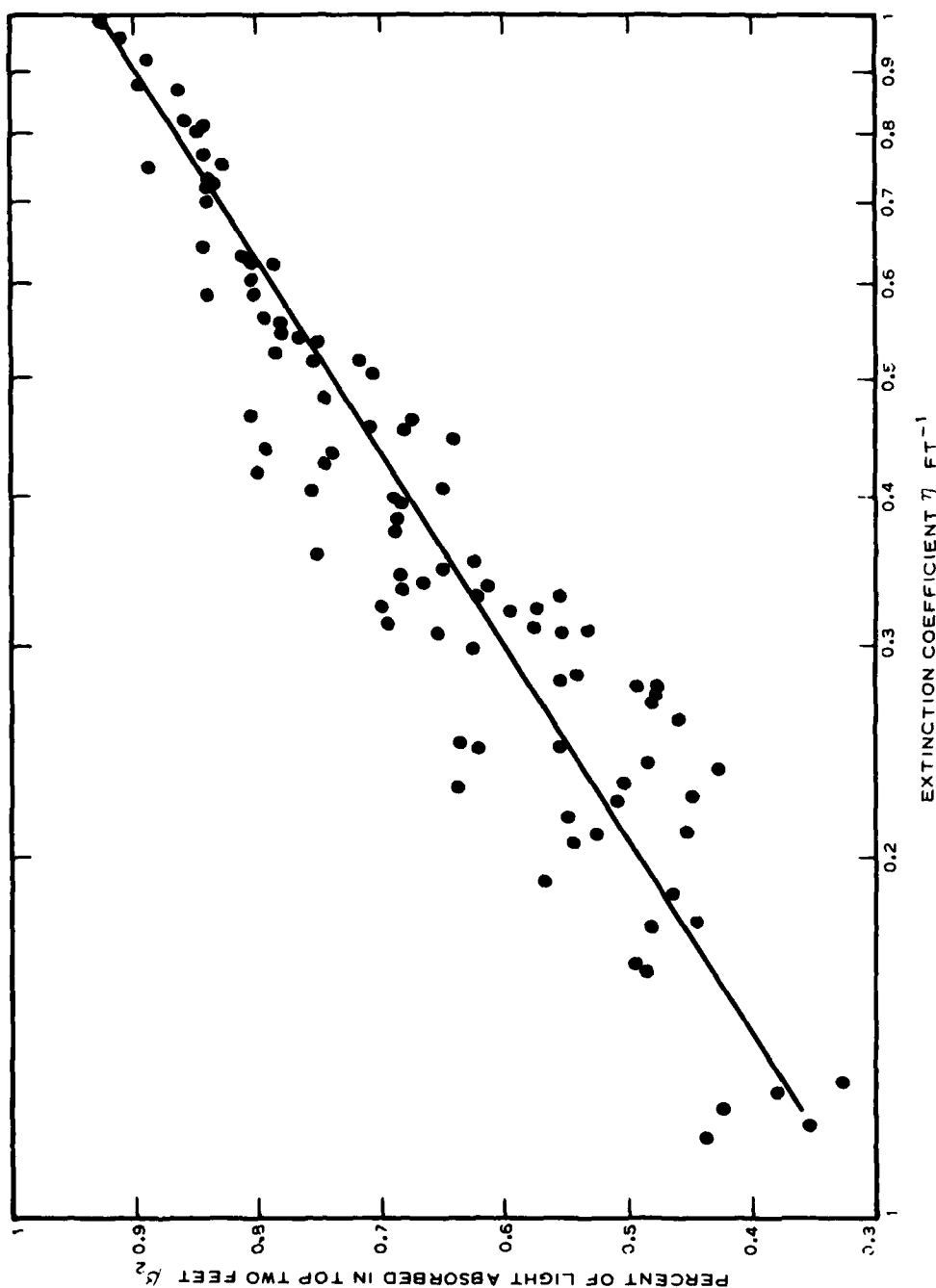


Figure 11. Relationship between the percent light absorbed in the top two feet and the extinction coefficient (G. Drummond, ORD, unpublished data).

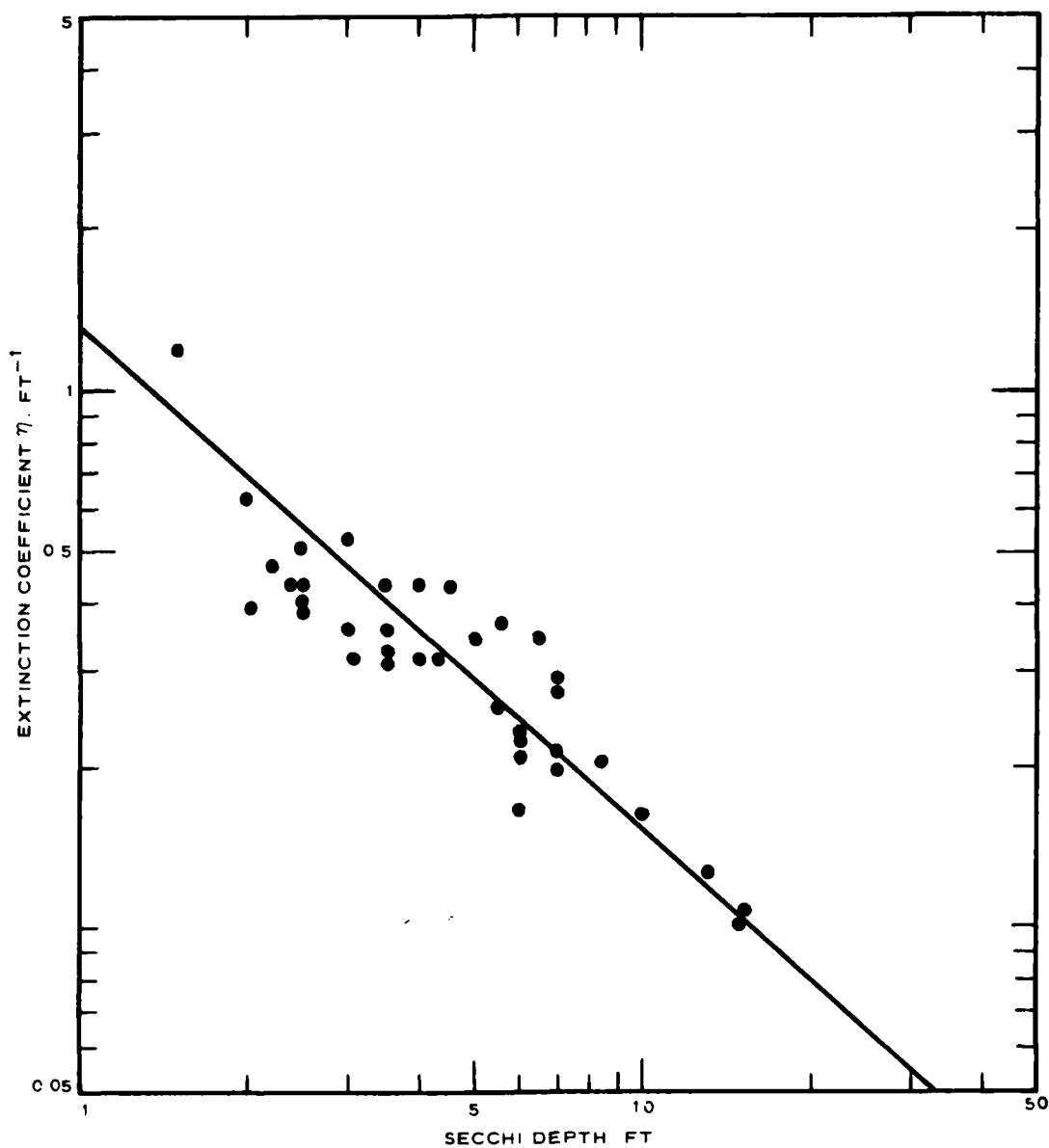


Figure 12. Relationship between the extinction coefficient and the Secchi disk depth (G. Drummond, ORD, unpublished data).

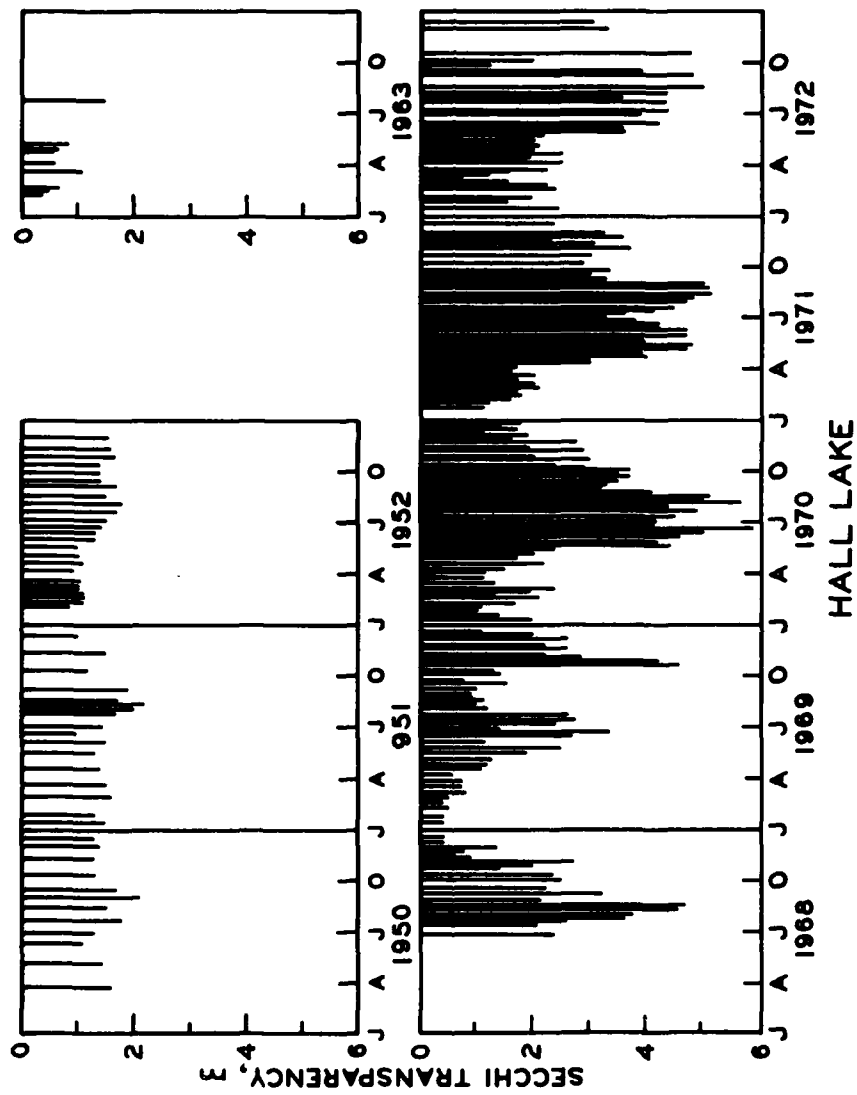


Figure 13. Typical seasonal variations in Secchi disk depths (after Culver, 1975).

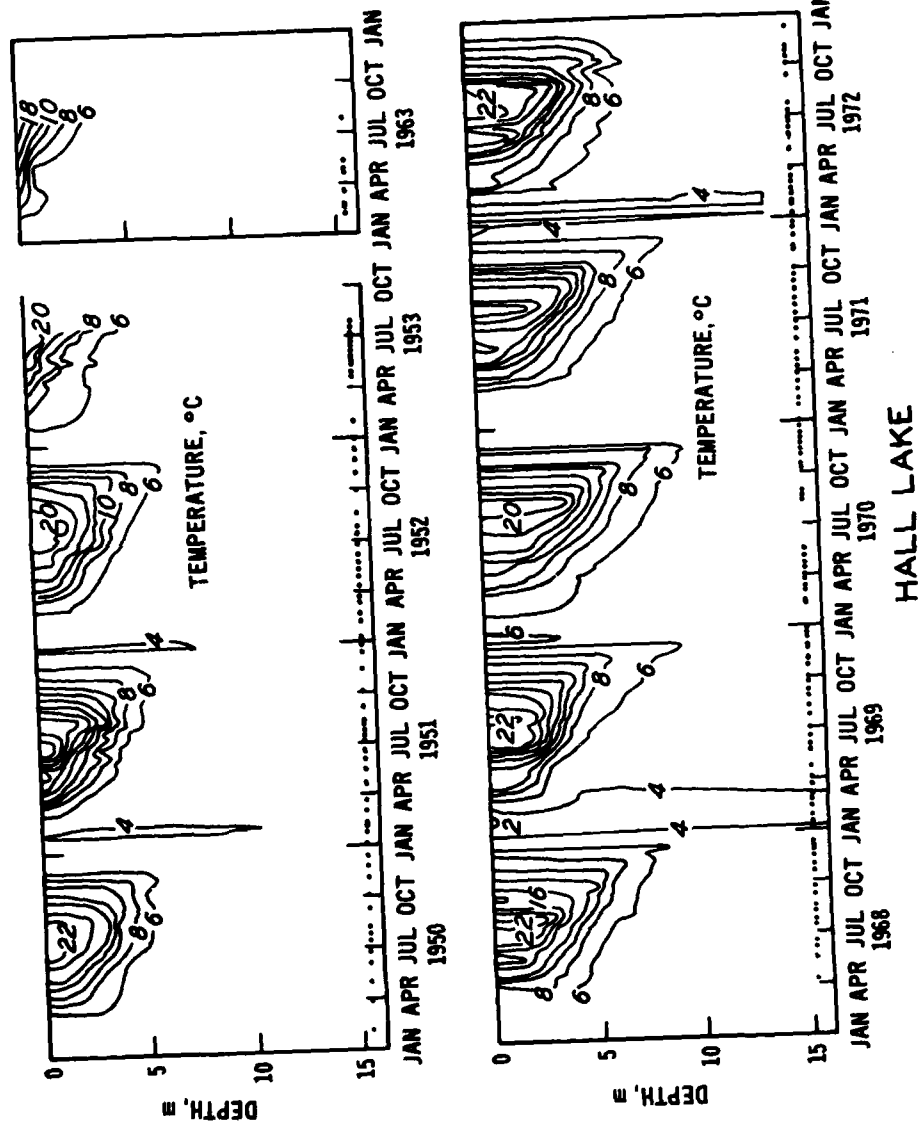


Figure 14. Effect of light penetration on the seasonal temperature cycle (after Culver, 1975). Compare with Figure 13.

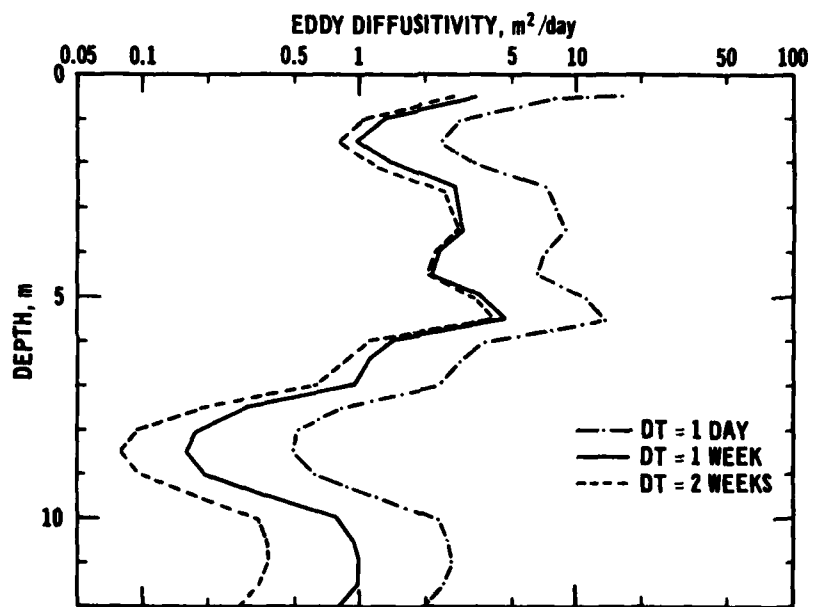


Figure 15. Spatial and temporal variations in back calculated eddy diffusivities, McCarrons Lake, Minnesota.

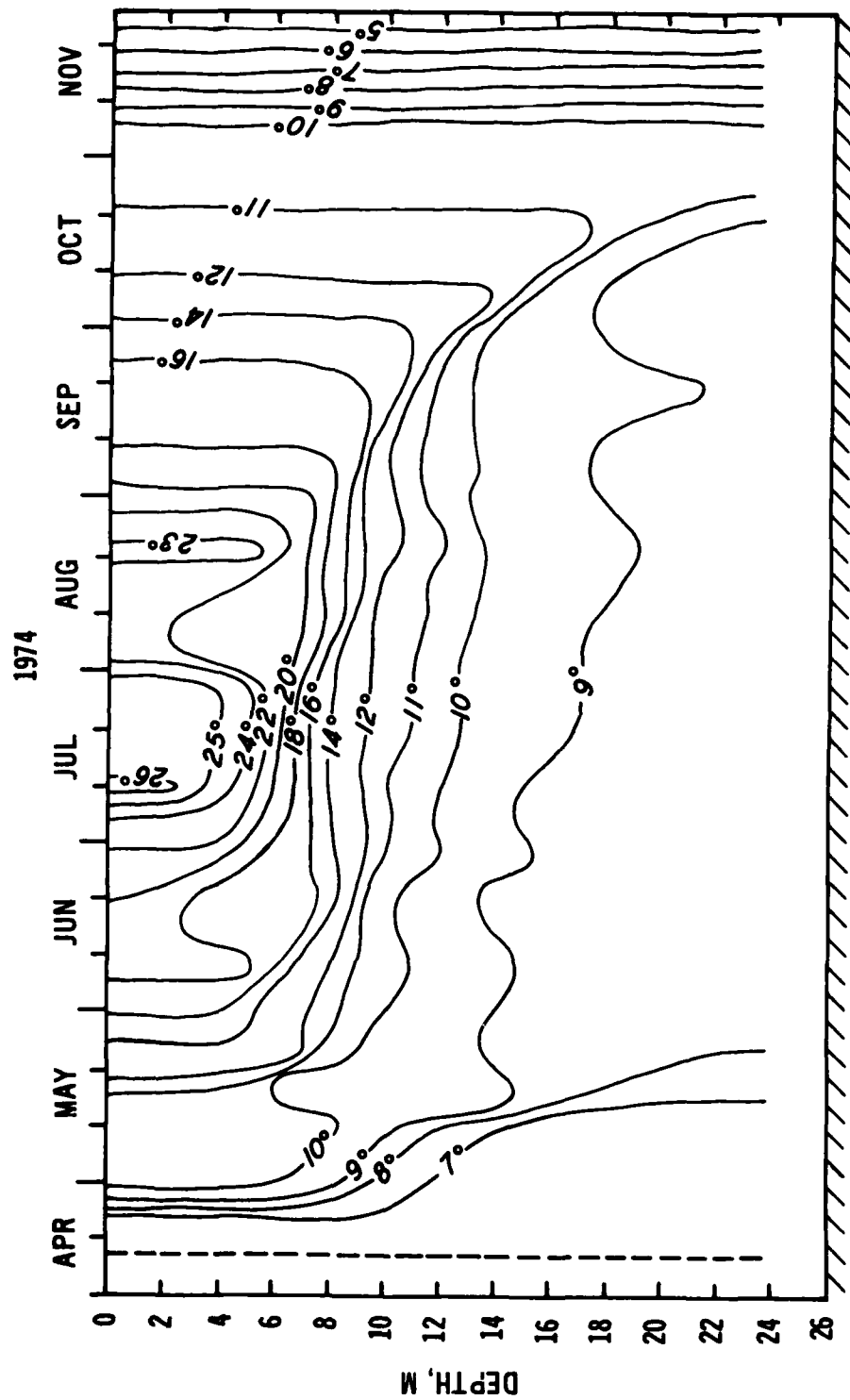


Figure 16. Typical seasonal temperature cycle for a dimictic lake, Lake Calhoun, Minnesota.

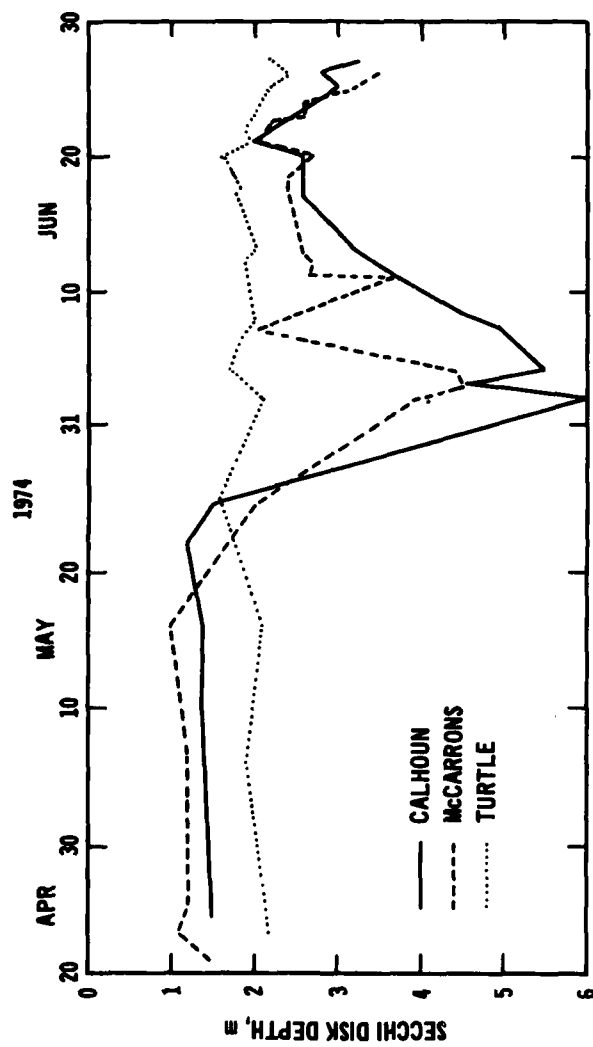
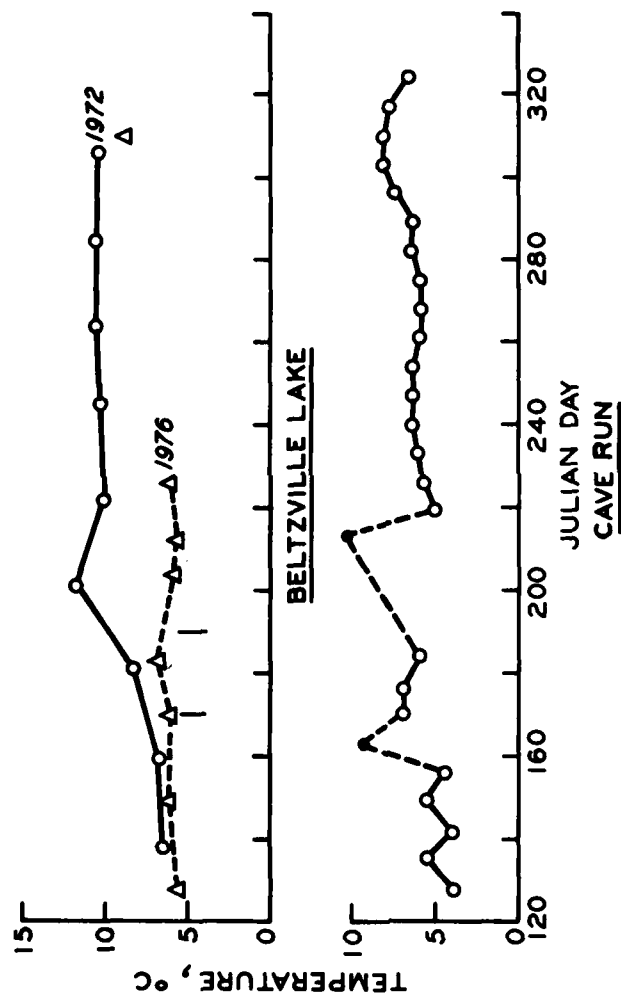


Figure 17. Temporal variations in Secchi disk depths.



HYPOLIMNION TEMPERATURE

Figure 18. Temporal variations of hypolimnion temperatures under stratified conditions.

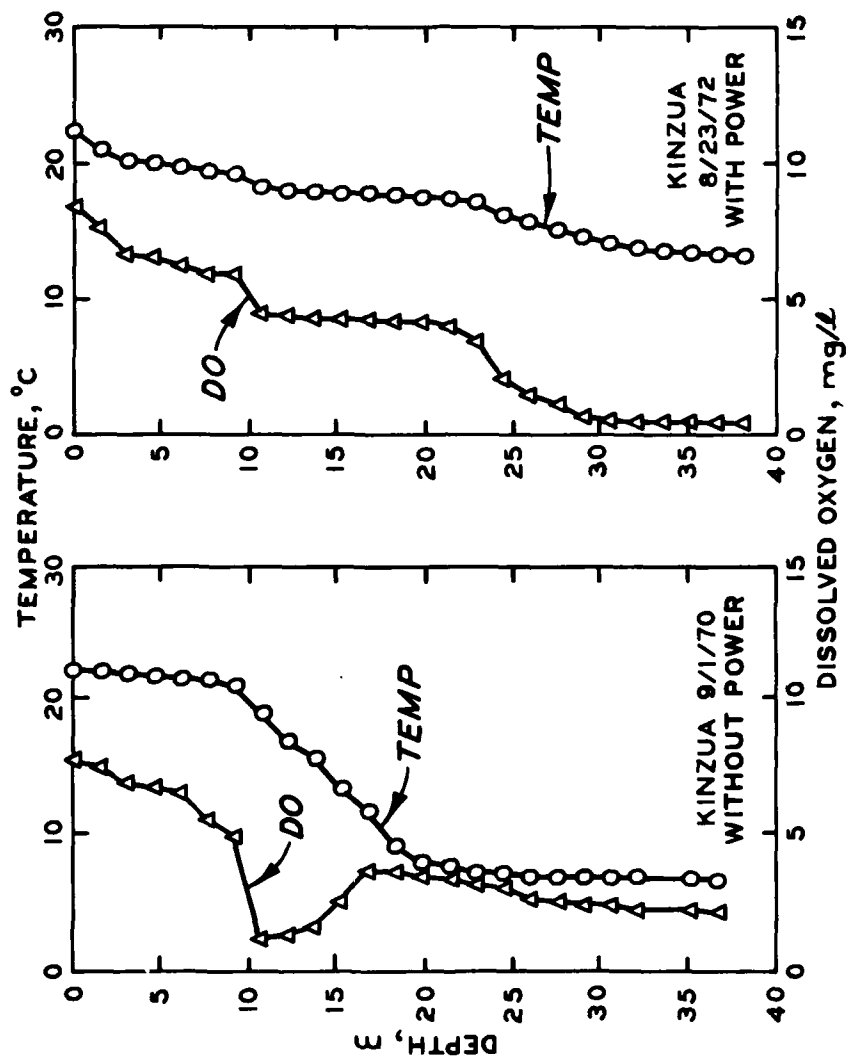


Figure 19. Effect of power operations on temperature and dissolved oxygen profiles.

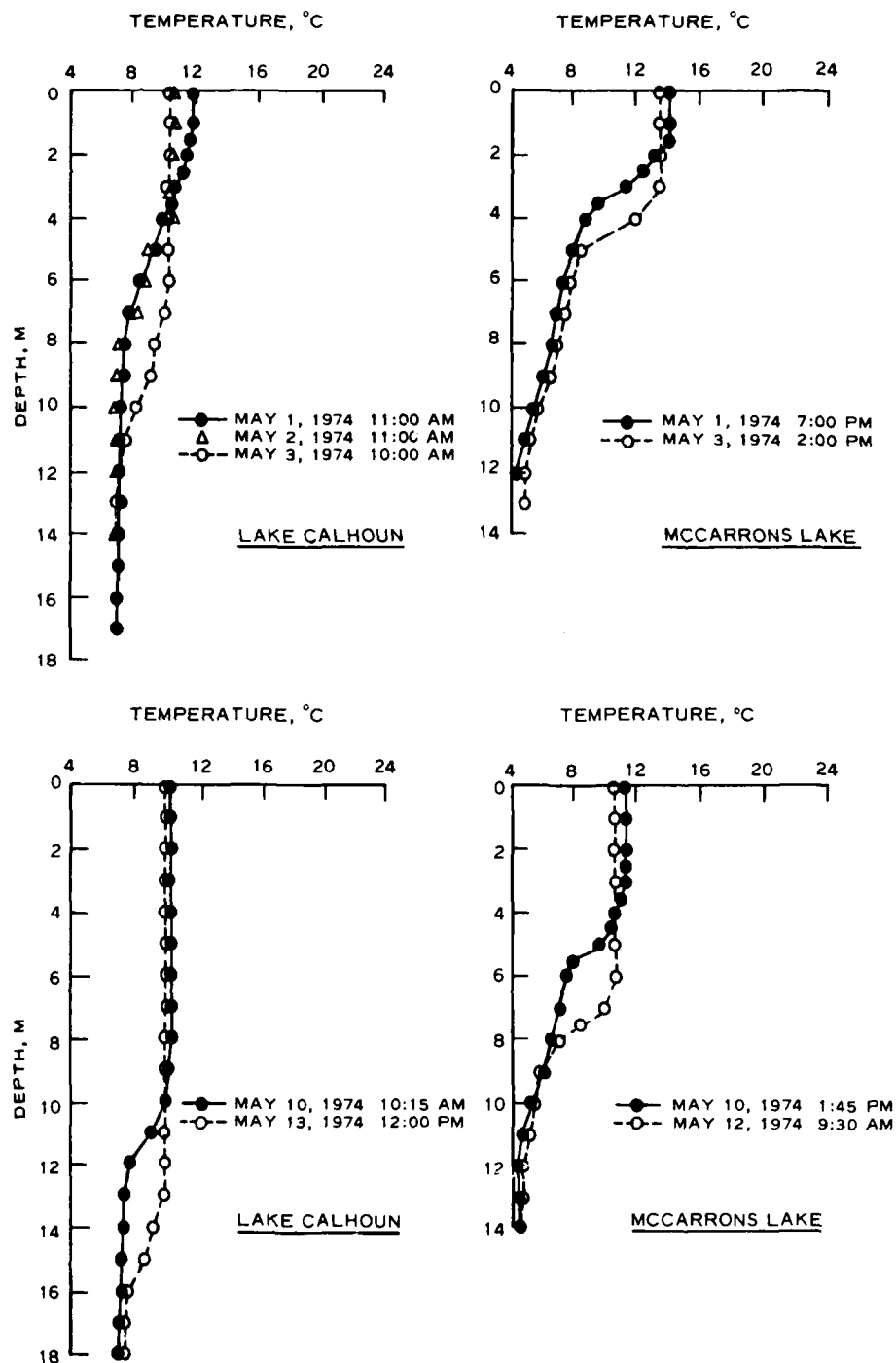
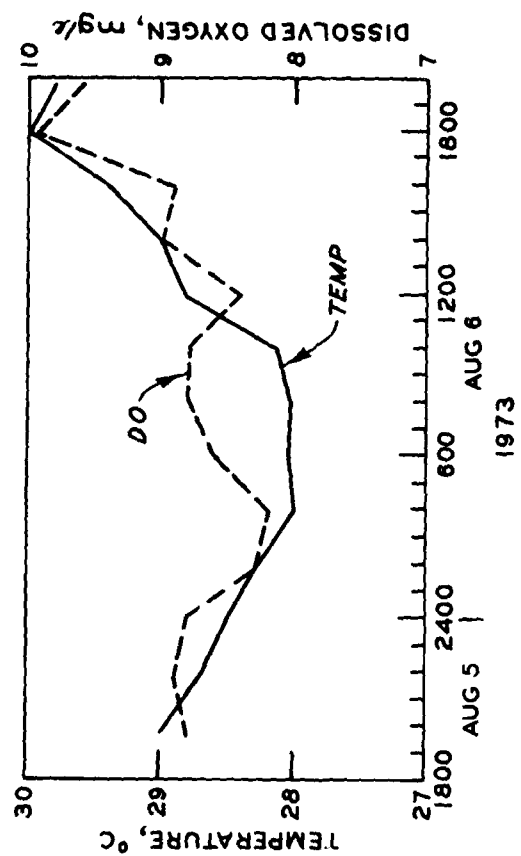


Figure 20. Examples of wind mixing.



DIURNAL - DEGRAY LAKE

Figure 21. Diel variations in temperature and dissolved oxygen.

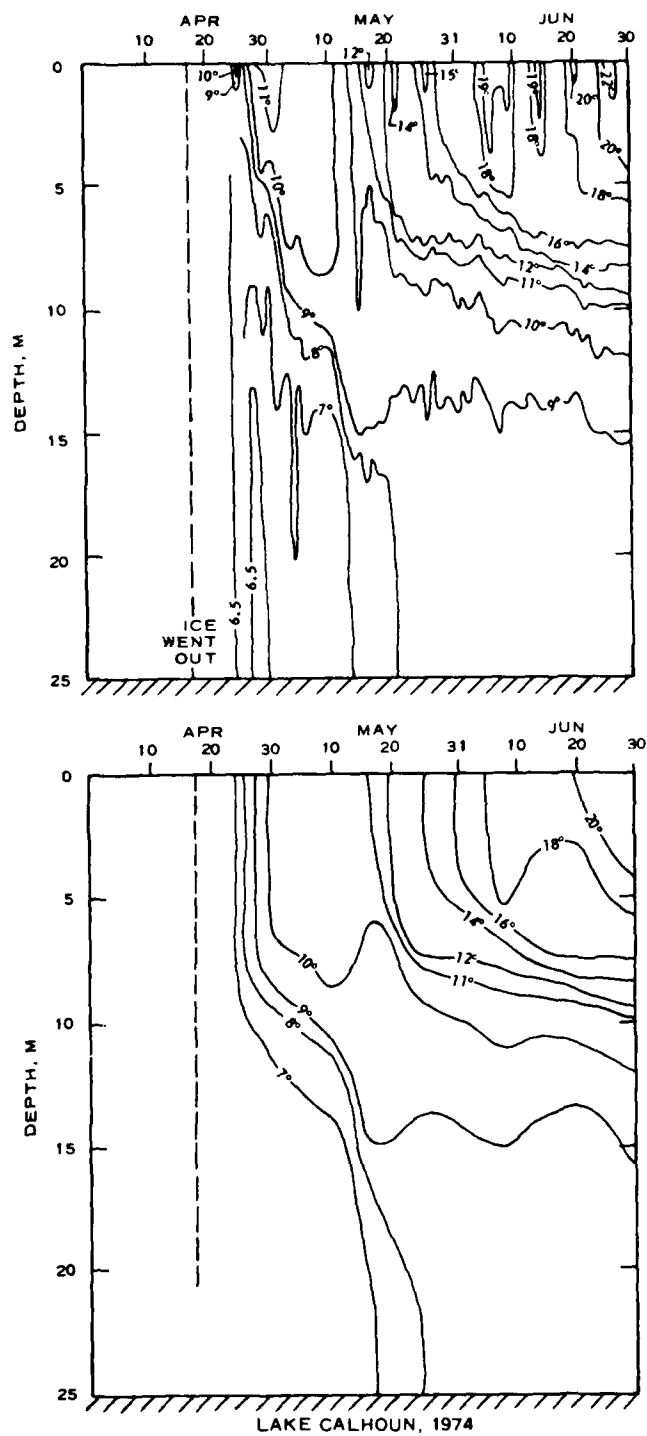


Figure 22. Comparison of daily (synoptic) and weekly isotherms.

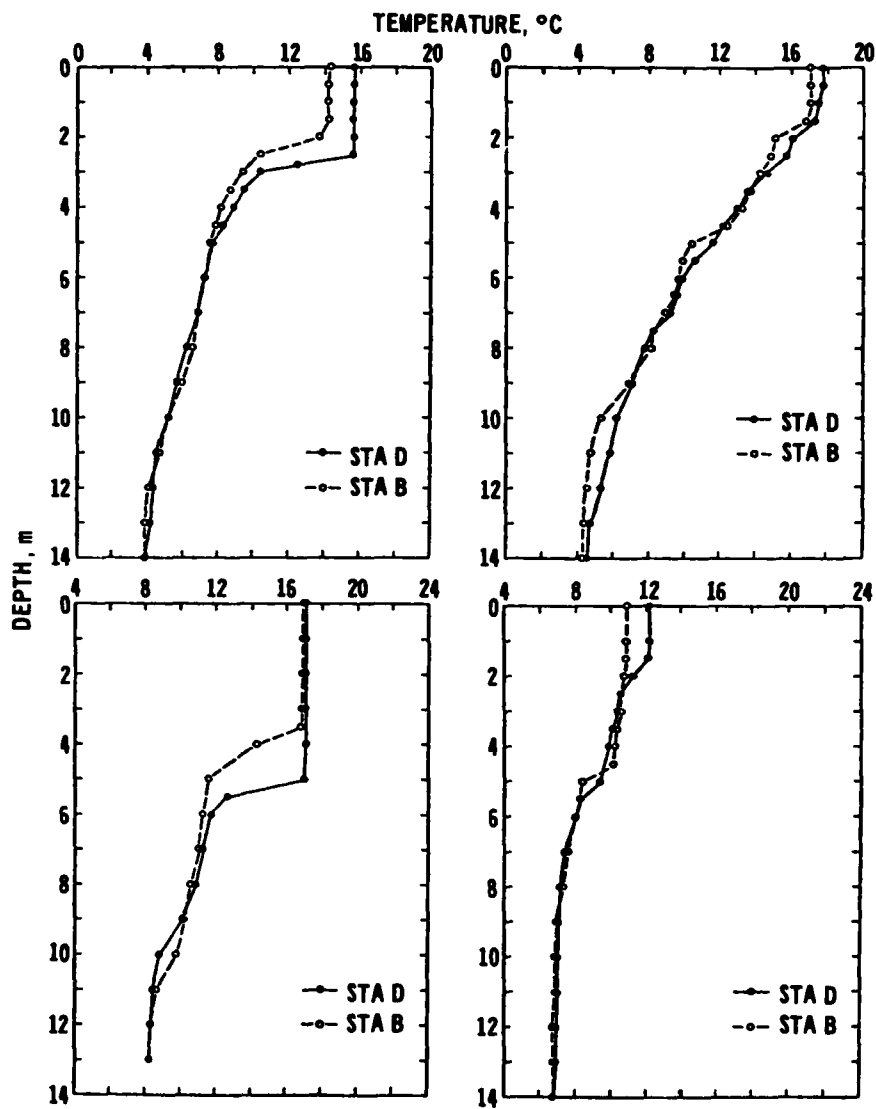


Figure 23. Horizontal variations in temperature profiles, McCarrons Lake, Minnesota.

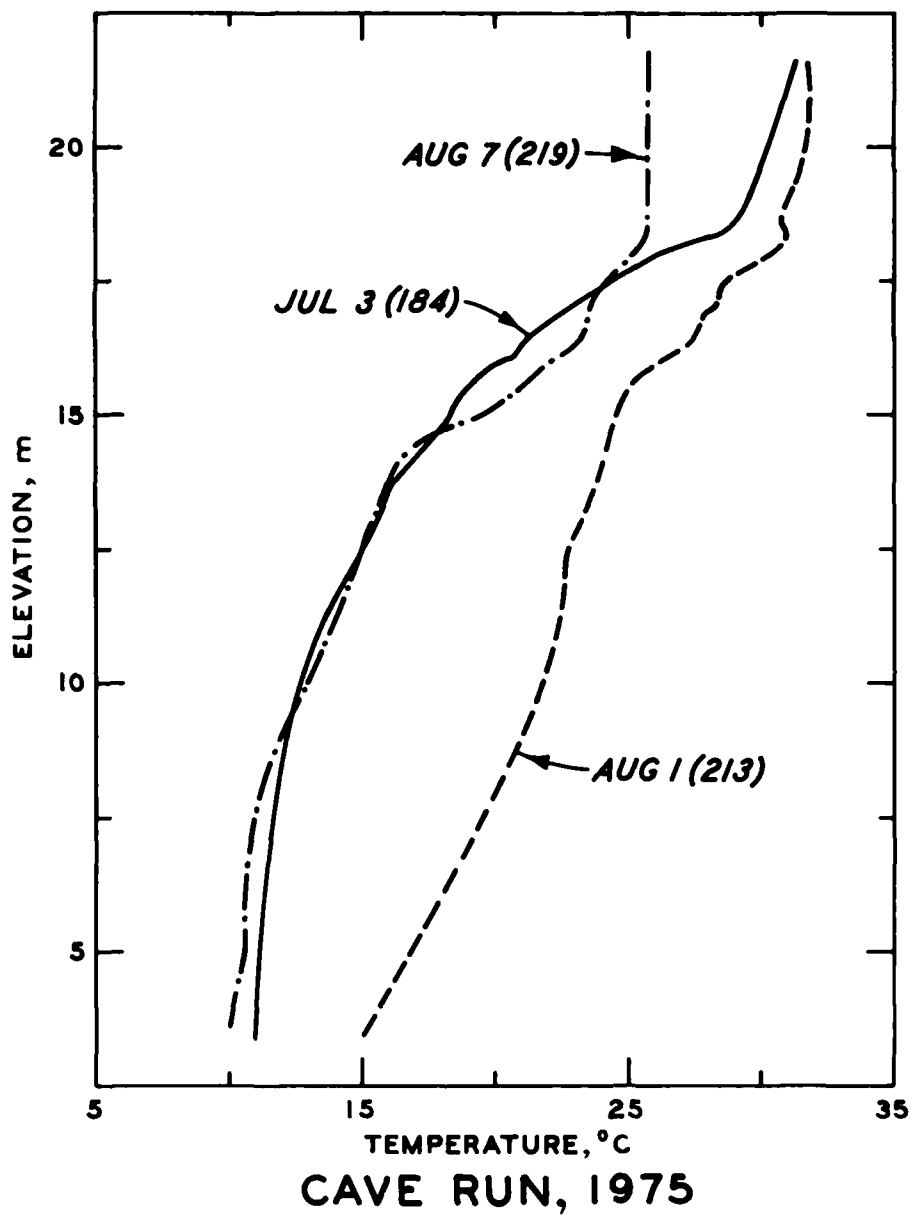


Figure 24. Example of erroneous temperature profile for day 213.

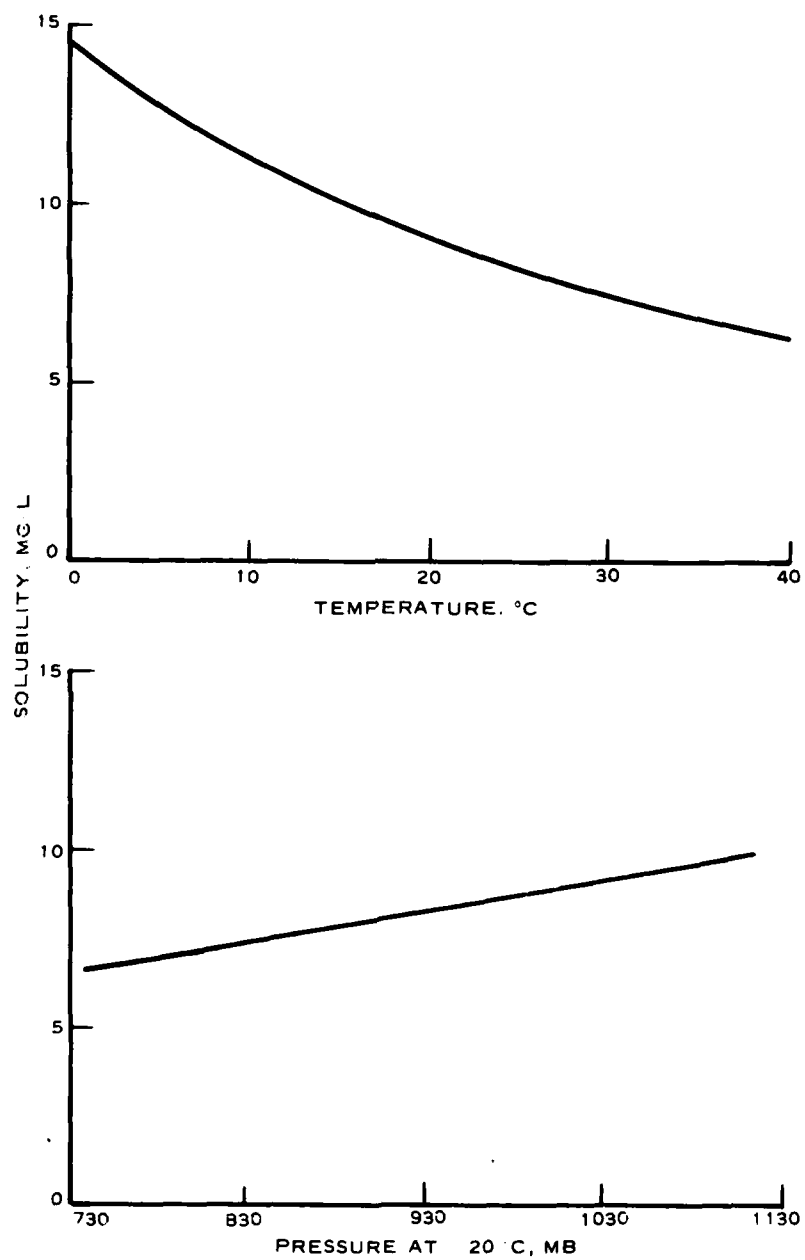
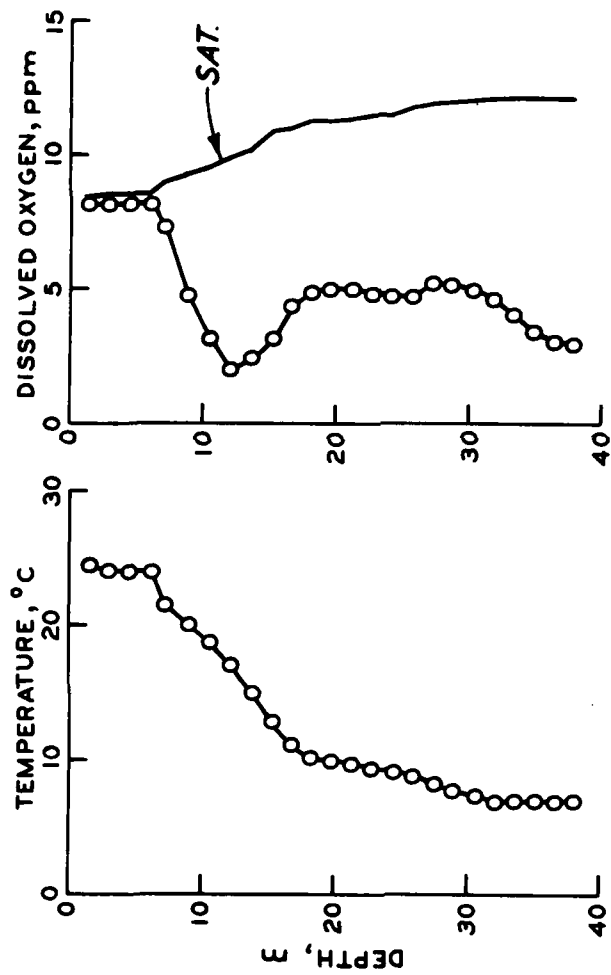
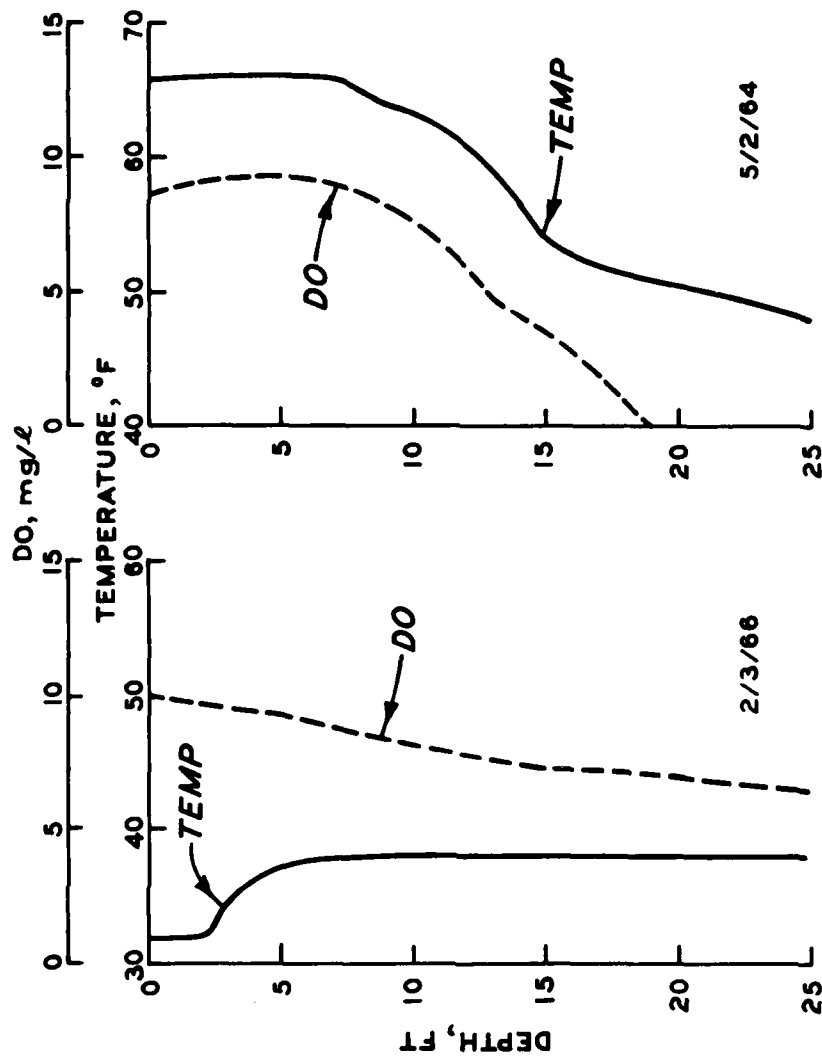


Figure 25. Solubility of oxygen in water.



BELTZVILLE LAKE, TEMPERATURE AND DISSOLVED OXYGEN PROFILES, 8/23/74

Figure 26. Comparison of measured and saturated DO profiles.



COX HOLLOW LAKE, WISCONSIN

Figure 27. Example of clinograde oxygen profile.

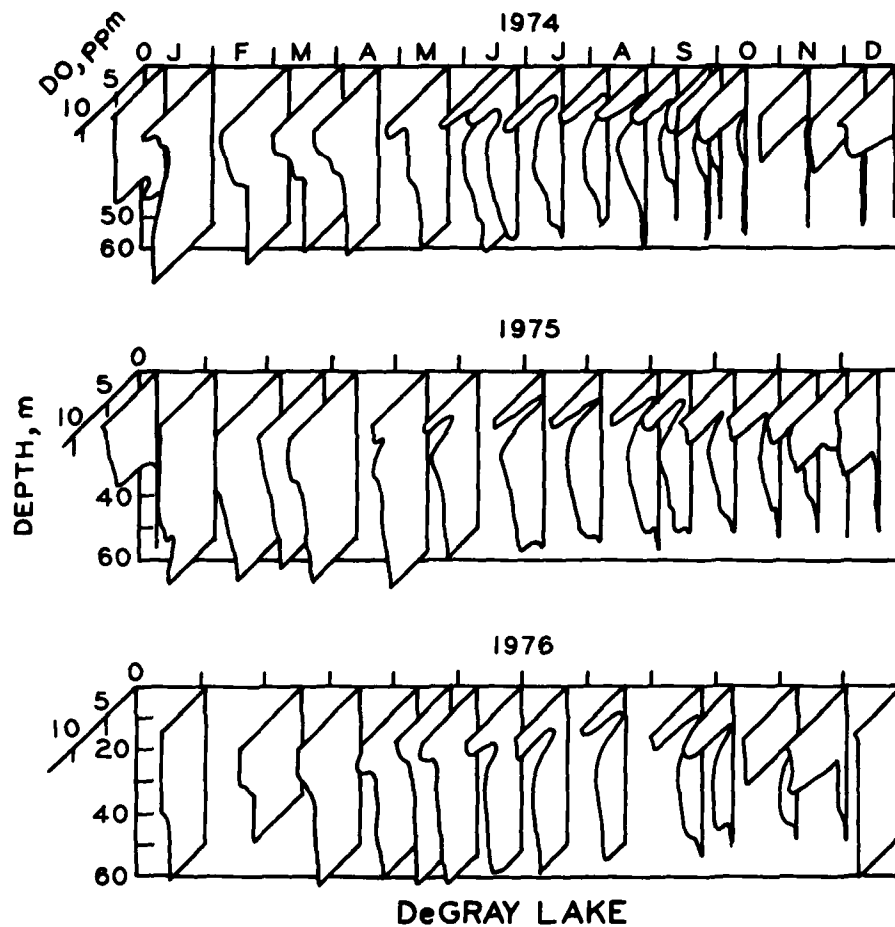


Figure 28. Seasonal variations in DO.

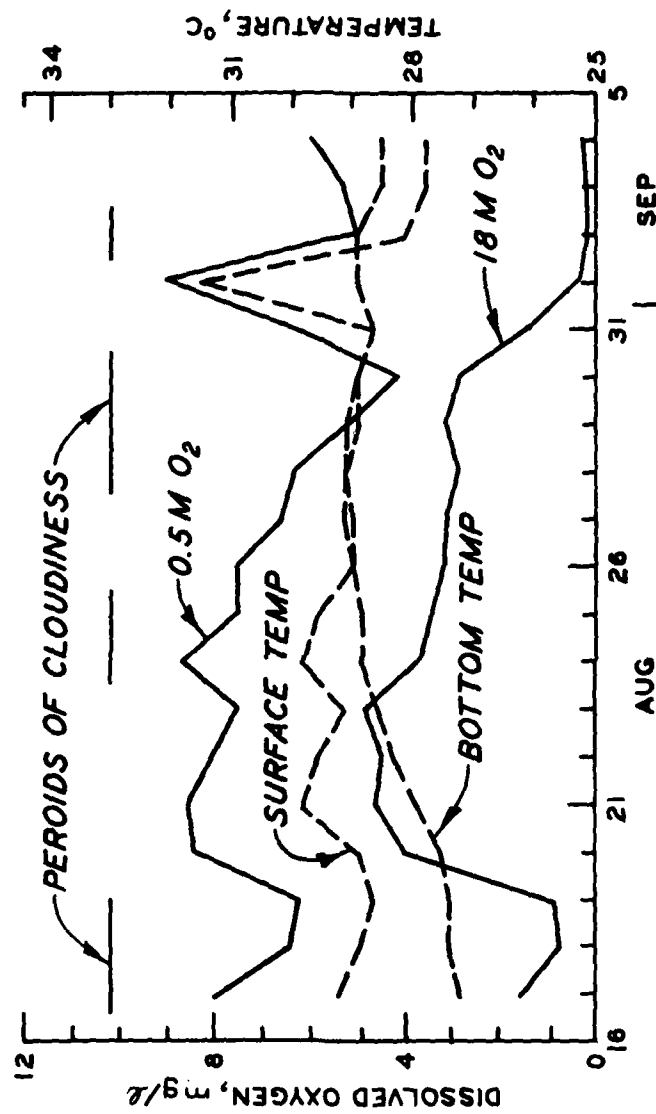


Figure 29. Synoptic variations in temperature and DO, Keystone Lake, OK (after Faust, 1972).

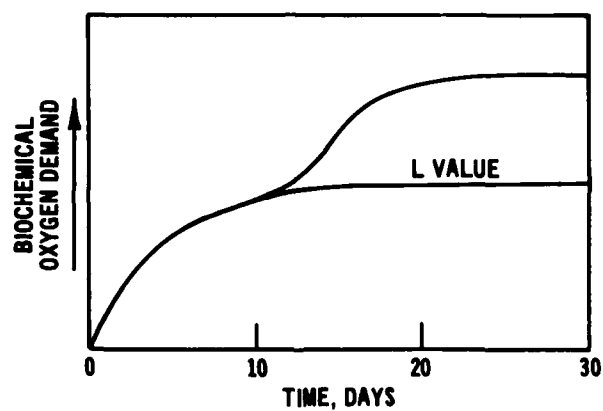


Figure 30. Classical BOD reaction where L is the 20-day BOD minus nitrification.

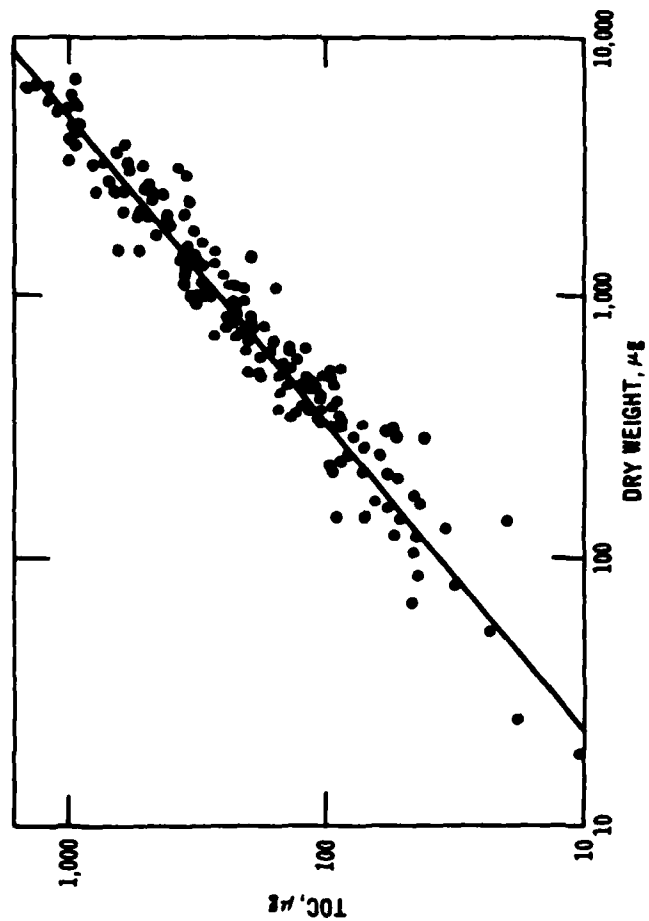


Figure 31. Relationship between TOC and zooplankton dry weight, DeGray Lake, Arkansas.

COLIFORMS AND WATER QUALITY:
USE OF DATA IN PROJECT DESIGN AND OPERATION

By

Kent W. Thornton*

INTRODUCTION

Coliform bacteria are important in many CE impoundments because of their impact on project purposes such as recreation and water supply. Knowledge of coliform transport and die-off may provide information for project design and operation that minimizes potential problems. The selection of appropriate recreation sites can be improved by knowing the path of coliform transport into and through an impoundment. This same information may be useful in determining withdrawal depths for water supply.

The coliform bacteria are defined to include all the aerobic and facultative anaerobic, gram-negative, nonsporulating bacilli that produce acid and gas from the fermentation of lactose (APHA, 1971). These organisms are primarily, but not exclusively, of human enteric or excretal origin. Two of the most common members of the coliform group are Escherichia coli and Enterobacter aerogenes.

Another group of bacteria of interest in a water quality sampling program are the fecal streptococci bacilli. Fecal streptococci are also found in the gastrointestinal tract of warm blooded animals, including man, but are typically nonpathogenic. Several species of fecal streptococci and their hosts include Streptococcus faecalis - man, S. bovis - cattle, and S. equinus - horses. Fecal streptococci are also found in hogs and fowl such as chickens, ducks, turkeys, etc.

The primary purpose of determining the coliform bacteria concentration in water is to indicate the possible presence of water-borne pathogenic organisms. Pathogenic organisms that cause typhoid, dysentery, and infectious hepatitis are also of human origin and may be transmitted through water. However, even during epidemics, these pathogenic organisms may be present in low concentrations, reducing the possibility of detection. In addition, culturing the organisms for identification is tedious and time consuming. For these reasons, a surrogate indicator such as fecal coliforms is desirable. To be an

* Research Ecologist, Environmental Laboratory, U. S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi 39180

effective indicator, a microorganism should possess several attributes: (1) it should be applicable to all types of waters such as lakes, streams, reservoirs, estuaries, etc.; (2) it should be present whenever pathogens are present with a survival time comparable to that of the hardest pathogen; and (3) it should not reproduce in contaminated waters thereby resulting in inflated values (Scarpino, 1974). Unfortunately, no indicator organism is known that satisfies all these criteria. Fecal coliform bacteria do, however, meet some of the requirements - they are of enteric origin, have similar die-off times as some pathogens, and have been correlated with an increase or decrease in some pathogens such as Salmonella typhi (typhoid). In addition, the coliform bacteria are relatively easy to culture.

This paper emphasizes the water quality relations of coliform distributions rather than public health aspects. While the two are inexorably related, the discussion centers on the interpretation of coliform data with respect to water quality. Sampling and culturing techniques and procedures are outlined in APHA Standard Methods (1975) and Millipore Application Manual AM302 (1973).

SOURCE IDENTIFICATION

While fecal coliform (FC) and fecal streptococci (FS) bacteria are present in both man and other livestock, their abundance varies from species to species (Table 1). The preponderance of bacteria in man consists of FCs. The reverse is true for livestock; that is, FS are the most abundant. This information can be used to assist in the identification of possible origins of waste input.

Since one group of organisms originate predominantly from man and the other group from livestock, a ratio of the two may provide information on the probable source of the waste discharge (Table 2). A FC to FS ratio greater than 4.0 would indicate human origin since, as indicated in Table 1, man is the only animal with a FC/FS ratio greater than 4.0. Similarly, the various species of livestock have FC/FS ratios less than 0.7. A range of FC/FS ratios from 2.0-4.0 or 0.7-1.0 indicates origins primarily from human or livestock origin, respectively. Ratios between 1.0 and 2.0 are inconclusive and are not really indicative of either source. In this instance, the sampling station should probably be moved upstream closer to the source of the waste discharge.

The above ratios were computed based on three underlying limitations: (1) the samples were collected no more than 24 hr travel time downstream from the source of the waste; (2) the pH of the water was within the range of 4.0-9.0; and (3) only FCs are compared with FS (Millipore, 1973). Differential die-off of the two groups after 24 hr obfuscates the conclusions that can be derived from the ratio values. This differential die-off minimizes the usefulness of the ratio. Extreme pH values also affect the survival of the bacterial groups,

again, resulting in differential die-off rates. Finally, consideration of total coliforms rather than FCs may result in erroneous conclusions since the proportion of total coliforms originating from human sources is highly variable. The total coliform group includes Enterobacter aerogenes that are ubiquitous in soil and may be washed off during storm events.

One use of coliform data, then, would be to determine the possible origin of the waste input and assess the seasonality of violations of the standard for body contact recreation (log mean of 200 FC cells/100 ml; EPA, 1976). Increased violations during the summer months, due to an influx of people may require additional package treatment plants at the project site or increased monitoring of inflows.

STORM INFLOWS

Total and fecal coliform concentrations generally increase in the inflow to a project during storm events. This increase is a function of many factors such as the size of the drainage area; magnitude of the storm event; land use in the basin; duration, intensity, and location of the rainfall; and other variables. In general, coliforms exhibit the "first-flush" phenomena. This implies the loading of coliforms to a stream occurs during the early stages of a storm. Perrier et al. (1977) studied the storm event loading of coliforms and other water quality variables to the Caddo River above DeGray Reservoir in south central Arkansas. Although there was a lot of variation in the data, the first-flush phenomena was evident (Figure 1).

Regression equations were derived to predict coliform loadings to the Caddo River as a function of streamflow during storm events (Perrier et al., 1977). Using data from the entire hydrograph to predict coliform concentrations resulted in an R^2 value of 0.19 or flow alone accounted for only 19 percent of the variance in the FC concentrations. Using data from only the rising limb of the hydrograph in the equation resulted in an R^2 value of 0.42 or flow alone now accounted for 42 percent of the variance in FC concentrations. The regression equation using data from the rising limb of the hydrograph

was

$$\log C = -0.00263 + 0.8177 \log Q \quad \text{S.E.} = 0.6736$$

where

C = FC concentration, cells/100 ml

Q = streamflow, m^3/s

S.E. = standard error

By monitoring several storm events a year for several years, it may be possible to quantify the relation between coliform concentrations and streamflow. This knowledge, coupled with the following discussion, may permit reservoir operations personnel to predict when coliform violations may occur and for how long they may persist. In addition, planners may be able to determine the most appropriate sites for recreation areas.

Given the magnitude of the streamflow and its associated coliform load, analyses can be performed to determine if coliform concentrations exceed standards and how far into the pool these violations may persist.

IN-POOL WATER QUALITY

Storm events have been monitored by the USAE Waterways Experiment Station (WES) in both the Caddo River and as the events proceeded through DeGray Reservoir (Figures 2-4). These storm events were monitored using a transmissometer to measure the change in the percent transmission. A percent transmission of 100 percent would indicate pure water, for example, whereas 0-percent transmission indicates very turbid water that permits no transmission of light through the light path. By using spatial and temporal variations in percent transmission, it was possible to trace the progress of stormflow into and through the reservoir. Knowledge of where the flow goes also permits an assessment of the fate of quality constituents.

The storm flow that occurred on 24 October 1976 entered as an underflow (Figure 5). As it proceeded into the reservoir, it appeared to become an interflow near Station 12 on 27 October 1976 (Figure 6). The inflowing streamwater had an initial temperature of ca. 15°C. A comparable density strata existed at around 11 or 12 m below the surface. Comparison with the temperature profile for this month (Figure 7) indicated 15°C temperature ca. 12 m below the surface. On 3 March, the inflowing stream waters were approximately 10°C while the reservoir was nearly isothermal at approximately 8 or 9°C (Figure 8). The streamflow initially followed the old stream channel as it entered the reservoir and stayed near the bottom (Figure 9). However, as the stormflow proceeded further into the reservoir, the buoyant forces of the slightly less dense water started to dominate and the flow rose in the pool until it moved primarily as an overflow. On 16 June, the inflowing stream water was approximately 19-20°C. It is evident from Figure 10 that the storm water proceeded into the pool as a rather narrow interflow at approximately 8 m. A review of Figure 11 will indicate a strong thermal stratification with the 20°C water located at approximately 8 m below the surface.

This information is quite useful in interpreting other water quality data. A perusal of data from Red Rock Reservoir (Figure 12)

will indicate a reasonable correspondence of suspended solids concentrations and FC concentrations. Coliforms may be washed into the stream unattached or adsorbed to solid particles. There is also evidence that some coliforms may enter the water column through bed scour (Grimes, 1975; Matson et al., 1978; Van Donsel and Geldreich, 1971). The use of percent transmission, then, may be an excellent means of tracking the initial movement of coliforms into the pool. If this is true, then there should be some correspondence between peak coliform concentrations in the pool and the expected movement of stormflows through the reservoir.

A comparison of coliform profiles at various sites in the reservoir following storm events demonstrated the correlation among coliforms and storm events. The storm event of 24 October 1976 entered initially as an underflow and then proceeded as an interflow at approximately the 10-m depth at Station 12. This same pattern is illustrated in the FC concentrations in the pool (Figure 13). Samples were taken at 1, 3, 5, 10 m and 5-m increments thereafter, so it is not possible to state the FC concentrations at intermediate depths that correspond to the depth of the maxima flow regime. By 27 October 1976, the inflow had proceeded to Station 12 in the reservoir. The peak concentrations were located around 10 m. Station 13 had a peak at 5 m, which corresponded roughly to the peak increase in turbidity on 25 October 1976. By 3 November 1976, the coliform concentration at both stations had diminished below the standard of 200 cells/100 ml due to both die-off and settling.

The storm event of 3 March 1977 provided additional insight into coliform transport. On 3 March the streamflow plunged into the pool following the thalweg. As it proceeded down the reservoir, the slightly warmer stream water began to rise toward the surface. On 3 March, the coliform concentration at the upper Station 13 was greatest near the bottom (Figure 14). On 5 March, the storm inflow had proceeded to Station 10 with the leading edge located at a depth of approximately 5 m. It is evident from Figure 14 that this same pattern occurred with respect to coliforms. The streamflow and associated FC concentration were distributed throughout more of the pool at Station 12. This was a function both of the settling of suspended solids and coliforms as well as the stormflow rising in the pool due to the buoyant density differences.

The settling process was demonstrated even more vividly by the storm event of 16 June 1977. The storm inflow was narrowly defined as an interflow at a depth of approximately 7-8 m on 18 June. The coliform concentration peaked sharply at 5 m on 18 June and diminished rapidly on either side at Station 12 (Figure 15). By 21 June, the inflow had proceeded to Station 8 (Figure 15). The FC concentrations were now distributed more evenly between 5 and 10 m at Station 12. With these flow patterns and coliform profiles, it is possible to compute information that is useful in locating recreational sites and in determining the possible impact of storm events on existing recreational sites or water supply.

APPLICATION OF STORM EVENT DATA

The information from the movement of the turbidity plumes can be used to compute estimates, albeit rough, of the speeds of various storm events as they proceed through the reservoir (Table 3). Since the distance among stations is known and the measurement times were recorded, it is possible to compute the speed in terms of miles per hour ($m/s \sim 0.5 \times mph$). In addition, recurrence intervals on a seasonal basis can be computed from long-term flow records to determine the frequency of occurrence of a given storm. In this example, based on climatological similarities, winter was considered December-February, spring represented March-May, and fall was represented by September-November. The seasonal basis is important since erroneous conclusions can be reached for operational purposes if an annual basis is used. The peak discharge for the 3 March 1977 storm was approximately 22,000 cfs while the peak discharge for the 16 June 1977 storm was ca. 13,000 cfs. Although the June storm had a lower peak discharge, on a seasonal basis, this was equivalent to a 1 in 100 yr event while the March storm was equivalent to a 1 in 44 yr event. If enough storm events are measured, it may be possible to roughly correlate the speed or change in speed with which a storm proceeds through a reservoir with the inflow discharge. While actual measurements are obviously the most desirable, it also may be possible to compute some of these speeds using infrared aerial or land-satellite photos of turbidity plume movement in a project, especially if the event is an overflow.

An assumption generally made for coliform die-off is that it is a first order process; that is, die-off is exponential. Gunnison (unpublished data) conducted a series of experiments to predict the effect of temperature on coliform die-off using natural lake water as the media rather than an artificial media. Gunnison (personal communication) obtained significantly greater die-off of coliforms using natural water than artificial media. The results of his study for three temperatures are shown in Table 4. Figure 16 was constructed for illustrative purposes using an initial concentration of 100,000 organisms and the equations of Gunnison. At 5°C, it takes approximately 15 days for the FC concentration to decrease from 100,000 cells/100 ml to 200 cells/100 ml. This time is decreased to ca. 4 days at 20°C with only a slight decrease in time (3.5 days) at 35°C. The three die-off rates may be plotted versus temperature and estimates of die-off rates at other temperatures interpolated (Figure 17). While the scientific validity of this interpolation is questionable, the values obtained appear to be reasonable. The die-off rate at 20°C is approximately double that at 10°C, which could be expected from a Q_{10} assumption of a doubling of a chemical reaction rate for each 10°C increase in temperature. Whether biological reaction rates obey the Q_{10} assumption is a point of contention, but it may be possible, then, to obtain estimates for various coliform die-off rates as a function of temperature under controlled conditions.

The storm event information on coliform concentrations collected in the field has already been presented. The peak FC concentrations have been tabulated as a function of reservoir station and storm event (Table 5). With several assumptions, this laboratory and field information can be used to assist in project operation and design.

If we assume there is an exponential die-off of coliforms as a function of temperature, then

$$C = C_0 e^{-xt}$$

where

C = coliform concentration at time t and some specified temperature

C_0 = initial coliform concentration

x = die-off rate as a function of temperature

t = time in days.

Then,

$$t = - \frac{\ln \left(\frac{C}{C_0} \right)}{x}$$

$$x = - \frac{\ln \left(\frac{C}{C_0} \right)}{t}$$

From the laboratory data, C_0 and x are known, t is an independent variable so C may be computed. From the field data, C and C_0 are known, and using t as the travel time, x may be computed, or x may be assumed to be identical to the laboratory rate at a given temperature and t computed.

Using this approach, it is then possible to compare field rates and times with laboratory rates and times. For example, if we assumed plug flow (Lagrangian reference frame) with no dilution or entrainment, etc., based on a given temperature and time of travel, we could compute a die-off rate. Conversely, if we compute or assume a die-off rate at a given temperature, we can compute the expected time required to reach the concentration of C cells/100 ml. Again, this assumes no dilution or losses due to settling. A comparison of these values is shown in Table 6.

For the storm event of 24 October 1976, there was excellent agreement between the two approaches indicating that, for this particular event, plug flow may have been a representative model. In general, however, the "actual" die-off and the time required were much greater and shorter, respectively, than predicted under laboratory conditions. This is expected due to the effects of dilution and settling in addition to die-off. The laboratory equations would, in general, therefore, represent a conservative case for coliform die-off and travel times.

This information can be used to assist in the location of recreational sites and in anticipating possible body contact standard violations (Table 7). By installing or using recording gages and recording thermographs on major tributaries into CE projects, both flow and temperature can be continuously monitored. Thermal model predictions and actual temperature profiles in the reservoir over several years provide good information on what the average temperature profile will be at a given time. Information from equations of flow versus FC concentrations provides estimates of the initial inflowing coliform concentrations. Using, first, the laboratory equations and appropriate rates, the time required to reach the 200 FC cells/100 ml can be computed. Once this time is obtained, information from a table such as Table 3 giving average speeds as a function of storm events can be used to compute how many miles into the reservoir the inflow will proceed before the coliform concentration decreases below the body contact standard. The seasonal frequency of the storm event can be used to assess the significance of this degree of penetration. For example, if a storm event that proceeds 6 miles into the project has a recurrence interval of 1 in 100 yr, the decision to locate sites further up the impoundment for greater recreational benefits may be reasonable.

This same approach may be used to determine when potential problems or violations may occur at an existing recreational site. This information may help determine where and when additional sampling and monitoring of existing sites is desirable to minimize health risks.

DISCUSSION AND LIMITATIONS

The decrease in coliform concentrations within the water column is, as suggested by Zanoni et al. (1978) perhaps, best described by a disappearance rate rather than die-off rate. Studies by Matson et al. (1978) indicate the sediments may serve as a reservoir for coliform bacteria. The effects of settling and entrainment result in a decrease in coliform concentrations but do not, necessarily, increase the actual die-off rate.

The magnitude of the sampling error and resolution in the field data must constantly be kept in mind. An example of this magnitude

in sampling error may be obtained by considering the average values of coliforms collected at discrete depths versus that of composite samples from DeGray Reservoir (Table 8). It is known that bacteria may exist in layers only a few centimeters thick and can easily be missed by sampling only at a fixed depth. The error in one instance is almost three orders of magnitude in estimating concentrations.

The site specificity of these conclusions is also important. The DeGray Reservoir has stable land use with minor point-source discharges. A system with changing land use and multiple, significant point sources will respond quite differently. The site specificity must be incorporated in all sampling programs as well as in any conclusions formed from the data set.

All conclusions and approaches are predicated on obtaining good field data from CE projects. The importance of good field data should never be underestimated. It is extremely difficult to refute sound conclusions drawn from a well-designed and implemented field data collection program.

REFERENCES

- American Public Health Association. 1971. Standard Methods for the Examination of Water and Wastewater. APHA. New York, 13th ed.
- _____ 1975, Ibid, 14th ed.
- Environmental Protection Agency. 1976. Quality Criteria for Water. U. S. Gov't Printing Office. Washington, D.C.
- Geldreich, E. E. and Kenner, B. A. 1969. Concepts of fecal streptococci in stream pollution. Jour. Wat. Pollut. Contr. Fed. 41: R336-352.
- Grimes, D. J. 1975. Release of sediment-bound fecal coliforms by dredging. Appl Microbiol. 29:109-119.
- Gunnison, D. 1977. Unpublished data. Environ. Lab. USAE Waterways Experiment Station, Vicksburg, Mississippi.
- Matson E. A., Hornor, S. G., and Buck, J. D. 1978. Pollution indicators and other microorganisms in river sediment. Jour. Wat. Pollut. Contr. Fed. 50:13-19.
- Millipore Corporation. 1973. Biological Analysis of Water and Wastewater. Application Manual AM302.
- Perrier, E. R., Westerdahl, H. E., and Nix, J. F. 1977. Water quality loadings during thirteen storms on the Caddo River, AR. Am. Soc. Agr. Engr. Meeting. Paper No. 77-2529.
- Scarpino, D. 1974. Human enteric viruses and bacteriophages as indicators of sewage pollution. Int. Symp. from Sea Outfalls. Pergamon Press.
- Van Donsel, D. J. and Geldreich, E. E. 1971. Relationships of Salmonellae to fecal coliforms in bottom sediments. Wat. Res. 5:1079-1091.
- Zanoni, A. E., Katz, W. J., Carter, H. H., and Whaley, R. C. 1978. An in situ determination of the disappearance of coliforms in Lake Michigan. Jour. Wat. Pollut. Contr. Fed. 50:321-330.

TABLE 1
AVERAGE INDICATOR*
10⁶ CELLS/g FECES

| ANIMAL | F COLI | F STREP | RATIO FC/FS |
|---------|--------|---------|----------------|
| MAN | 13.0 | 3.0 | 4.4 |
| DUCK | 33.0 | 54.0 | 0.6 |
| SHEEP | 16.0 | 38.0 | 0.4 |
| CHICKEN | 1.3 | 3.4 | 0.4 |
| COW | 0.2 | 1.3 | 0.2 |
| PIG | 3.3 | 84.0 | 0.04 |

* GELDREICH & KENNER, 1969.

TABLE 2

INTERPRETATION
FC/FS RATIO

| | |
|-----------------------|----------------------|
| ● $FC/FS > 4.0$ | HUMAN WASTE |
| ● $FC/FS < 0.7$ | LIVESTOCK WASTE |
| ● $2 < FC/FS < 4.0$ | PRIMARILY HUMAN |
| ● $0.7 < FC/FS < 1.0$ | PRIMARILY LIVESTOCK |
| ● $1.0 < FC/FS < 2.0$ | ? MOVE NEARER SOURCE |

TABLE 3
STORM EVENT MOVEMENT

| STORM EVENT | ΔT (DAY) | MILES | MPH | INIT STA | TERM STA | RECURRENCE INTERVAL (YR) |
|----------------|---------------------|-------|------|-------------|-------------|-----------------------------|
| 10/24/76 | 1.12 | 8.9 | 0.33 | 16 | 13 | 1 IN 25 |
| 3/3/77 | 0.58 | 10.3 | 0.74 | 16 | 13,12 | 1 IN 44 |
| 3/5/77 | 1.91 | 6.1 | 0.13 | 13,12 | 10 | |
| 6/17/77 | 1.41 | 8.9 | 0.26 | 16 | 13 | 1 IN 98 |
| 6/18/77 | 1.12 | 8.6 | 0.32 | 13 | 10 | |

TABLE 4
EQUATIONS DESCRIBING
COLIFORM DIE-OFF

| | | |
|--|-------|--------------|
| $Y = 3.1 \times 10^6 \text{ EXP}(-0.42T)$ | 5° C | $R^2 = 0.84$ |
| $Y = 11.4 \times 10^6 \text{ EXP}(-1.64T)$ | 20° C | $R^2 = 0.99$ |
| $Y = 5.7 \times 10^6 \text{ EXP}(-1.86T)$ | 35° C | $R^2 = 0.97$ |

TABLE 5

STORM EVENT DATA

| STORM EVENT | STATION | | |
|----------------|---------|------|-----|
| | 16 | 13 | 10 |
| 10/24/76 | 14,600* | 3600 | |
| 3/3/77 | 5,700 | 1250 | |
| 3/5/77 | | 1250 | 590 |
| 6/17/77 | 60,000 | 560 | |
| 6/18/77 | | 560 | 26 |

* F COLI CELLS/100 ML.

TABLE 6

COMPARISON OF COLIFORM DIE-OFF RATES AND TIMES

| STORM EVENT | TEMP °C | RATE (DAY ⁻¹) | | TIME (DAY) | |
|----------------|------------|---------------------------|--------|------------|--------|
| | | LAGRANGIAN | ACTUAL | THEOR | ACTUAL |
| 10/24/76 | 15 | -1.23 | -1.25 | 1.13 | 1.12 |
| 3/3/77 | 10 | -0.82 | -2.61 | 1.85 | 0.58 |
| 3/5/77 | 10 | -0.82 | -0.39 | 0.91 | 1.91 |
| 6/17/77 | 20 | -1.64 | -3.31 | 2.85 | 1.41 |
| 6/18/77 | 20 | -1.64 | -2.74 | 1.87 | 1.12 |

TABLE 7

COMPUTATION

GIVEN: STORM DISCHARGE, Q
TEMPERATURE

INITIAL CONC

$$\text{LOG } C = 0.00263 + 0.8177 \text{ LOG } Q$$

TIME CALCULATION

$$T = -\frac{\ln\left(\frac{Y}{A}\right)}{X}$$

Y = 200 CELLS/100 ML

A = C

X = LAB OR KNOWN FROM FIELD CALC

DISTANCE

$$\text{MILES} = \text{MPH} \times \text{TIME}$$

TABLE 8

STATION 12

| <u>DATE</u> <u>1977</u> | <u>MEAN FOR DEPTHS</u> <u>0, 3, 5 METERS</u> <u>TOTAL COLIFORMS</u> <u>CELLS/100 ML</u> | <u>COMPOSITE OF</u> <u>UPPER 5 METERS</u> <u>TOTAL COLIFORMS</u> <u>CELLS/100 ML</u> |
|----------------------------|--|---|
| 5-27 | 453 | 1420 |
| 5-31 | 1967 | 2000 |
| 6-2 | 367 | 31000 |
| 6-6 | 633 | 90000 |
| 6-9 | 1600 | 85000 |
| 6-12 | 533 | 900 |
| 6-15 | 5000 | 73000 |
| 6-16 | 1033 | 26000 |
| 6-18 | 1333 | 51000 |
| 6-21 | 1500 | 20000 |
| 6-23 | 1800 | 28700 |

FIGURE 1

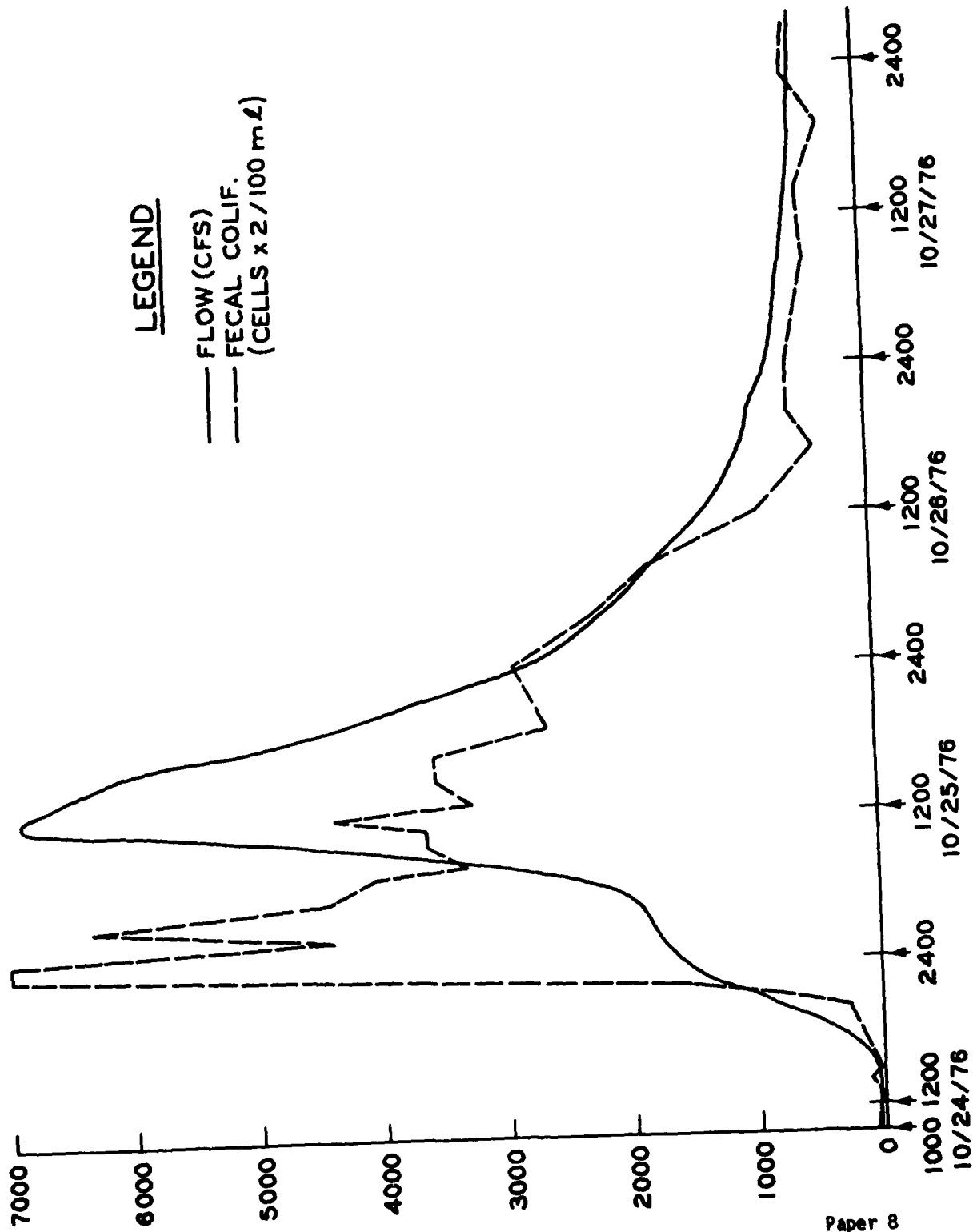
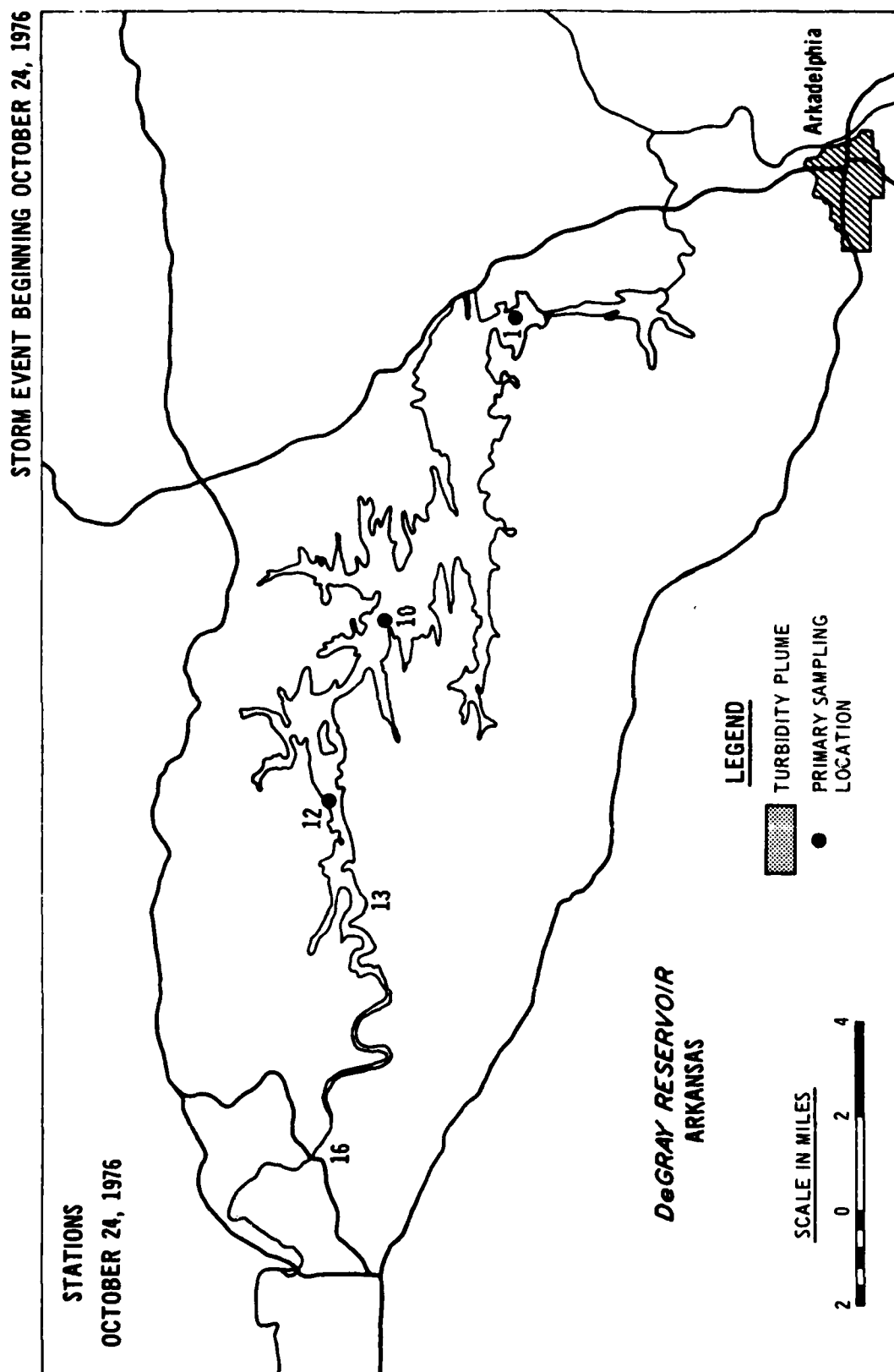


FIGURE 2a



STORM EVENT BEGINNING OCTOBER 24, 1976

SURFACE OBSERVATION - MUD PLUME
15:00 HR OCTOBER 25, 1976

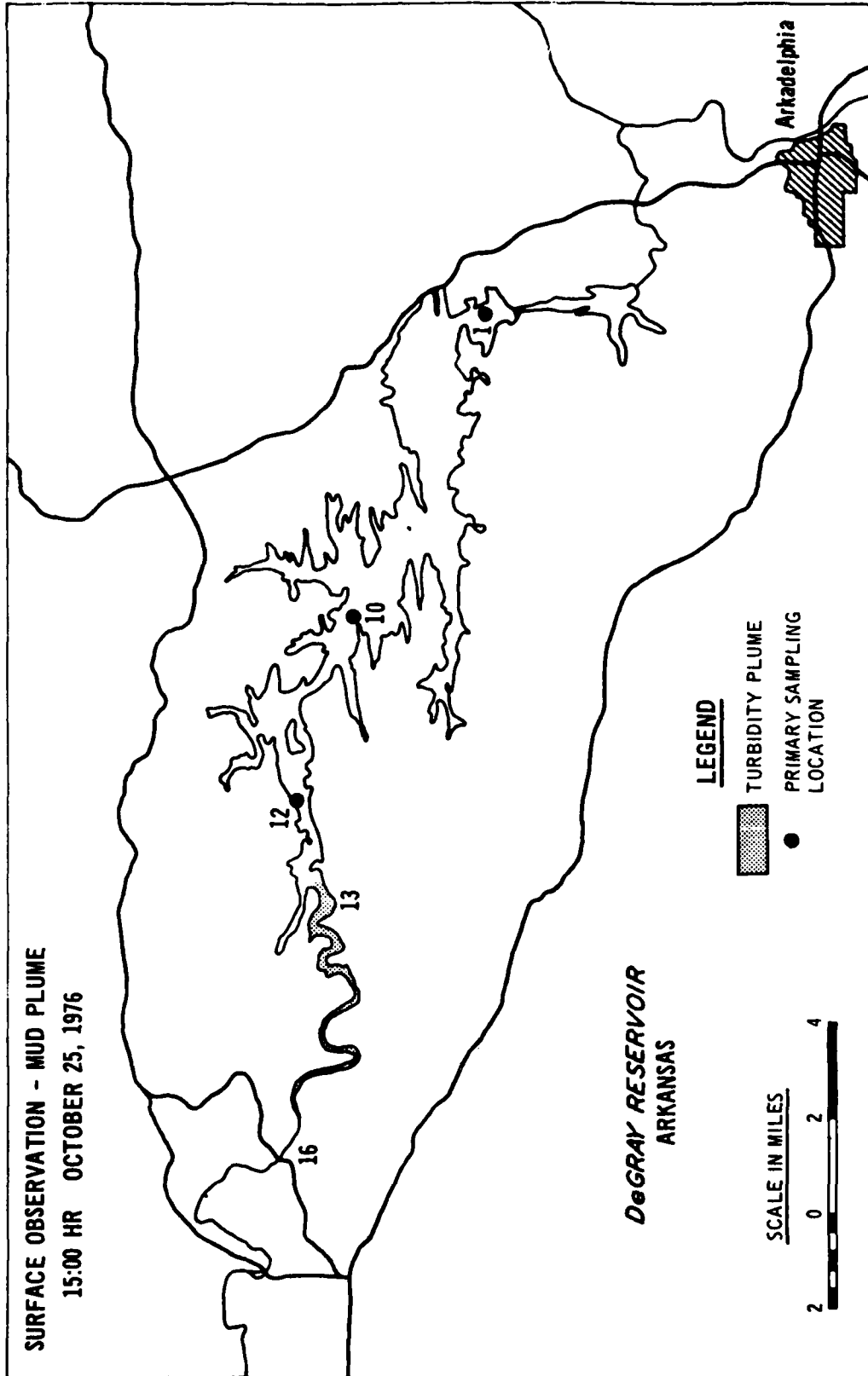
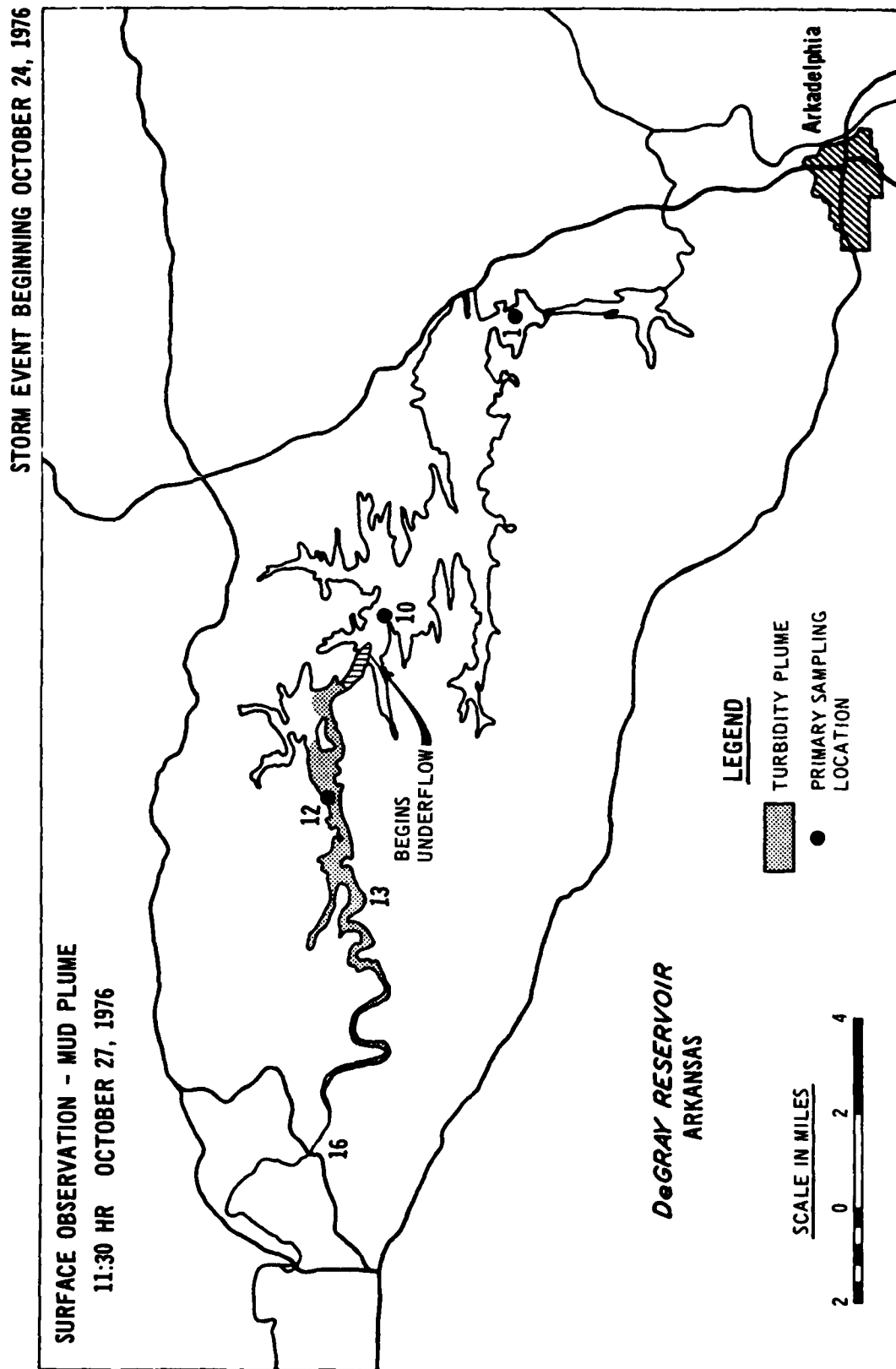


FIGURE 2b

FIGURE 2c



STORM EVENT BEGINNING MARCH 3, 1977

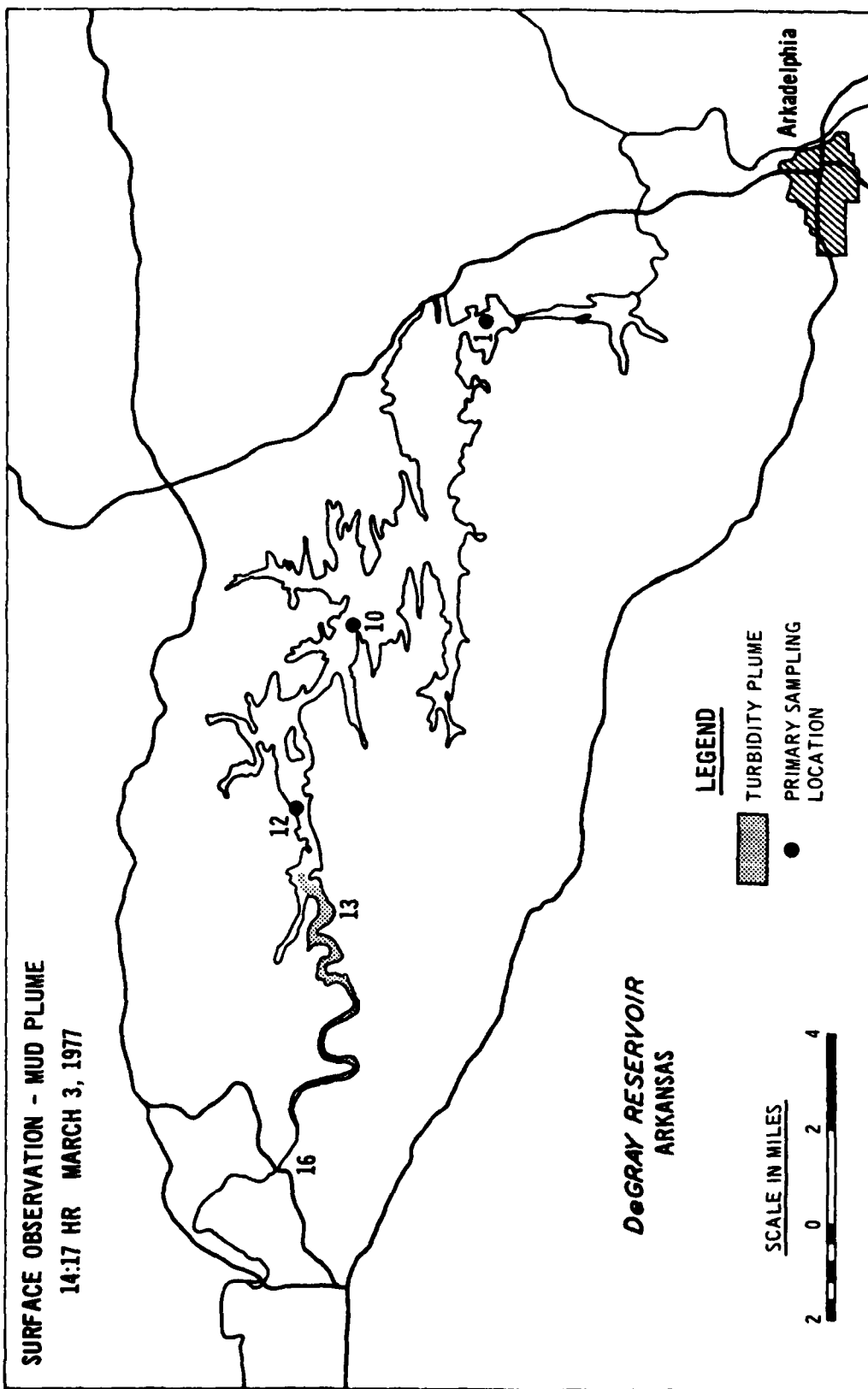


FIGURE 3a

STORM EVENT BEGINNING MARCH 3, 1977

SURFACE - PERCENT TIME OBSERVATION - MUD PLUME
12:40 HR MARCH 5, 1977

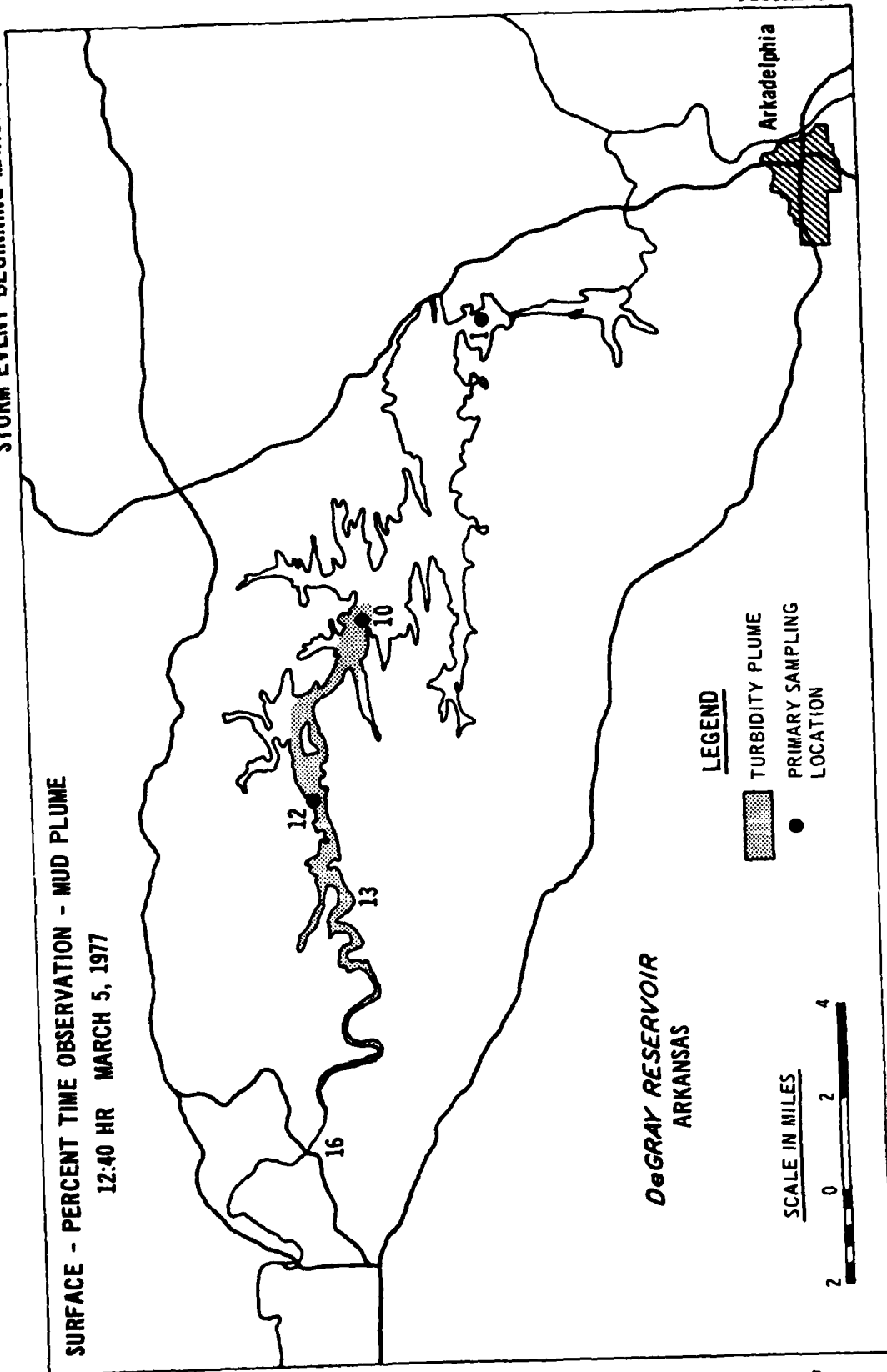
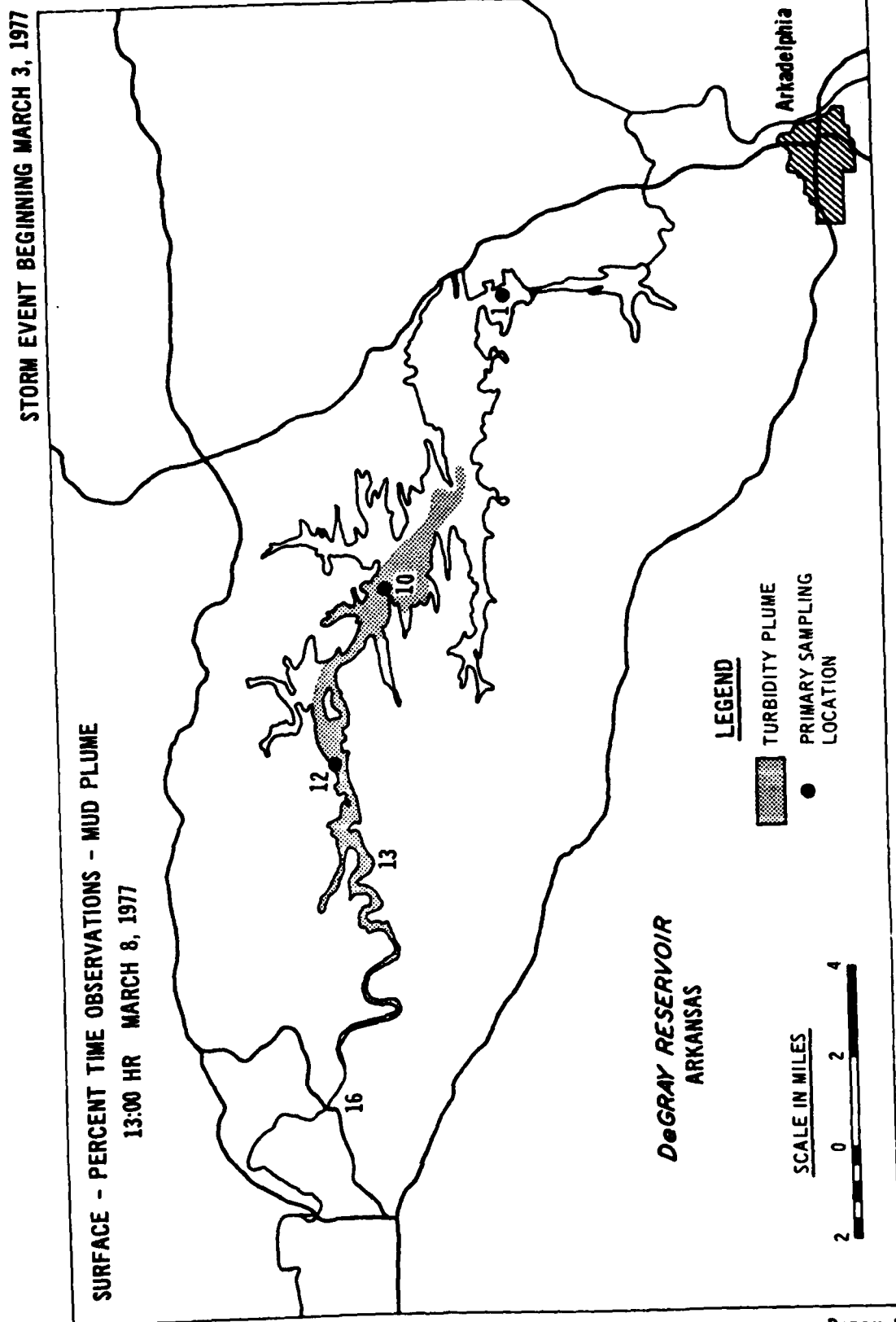


FIGURE 3b

FIGURE 3c



STORM EVENT BEGINNING MARCH 3, 1977

SURFACE - PERCENT TIME OBSERVATIONS
12:00 HR MARCH 9, 1977

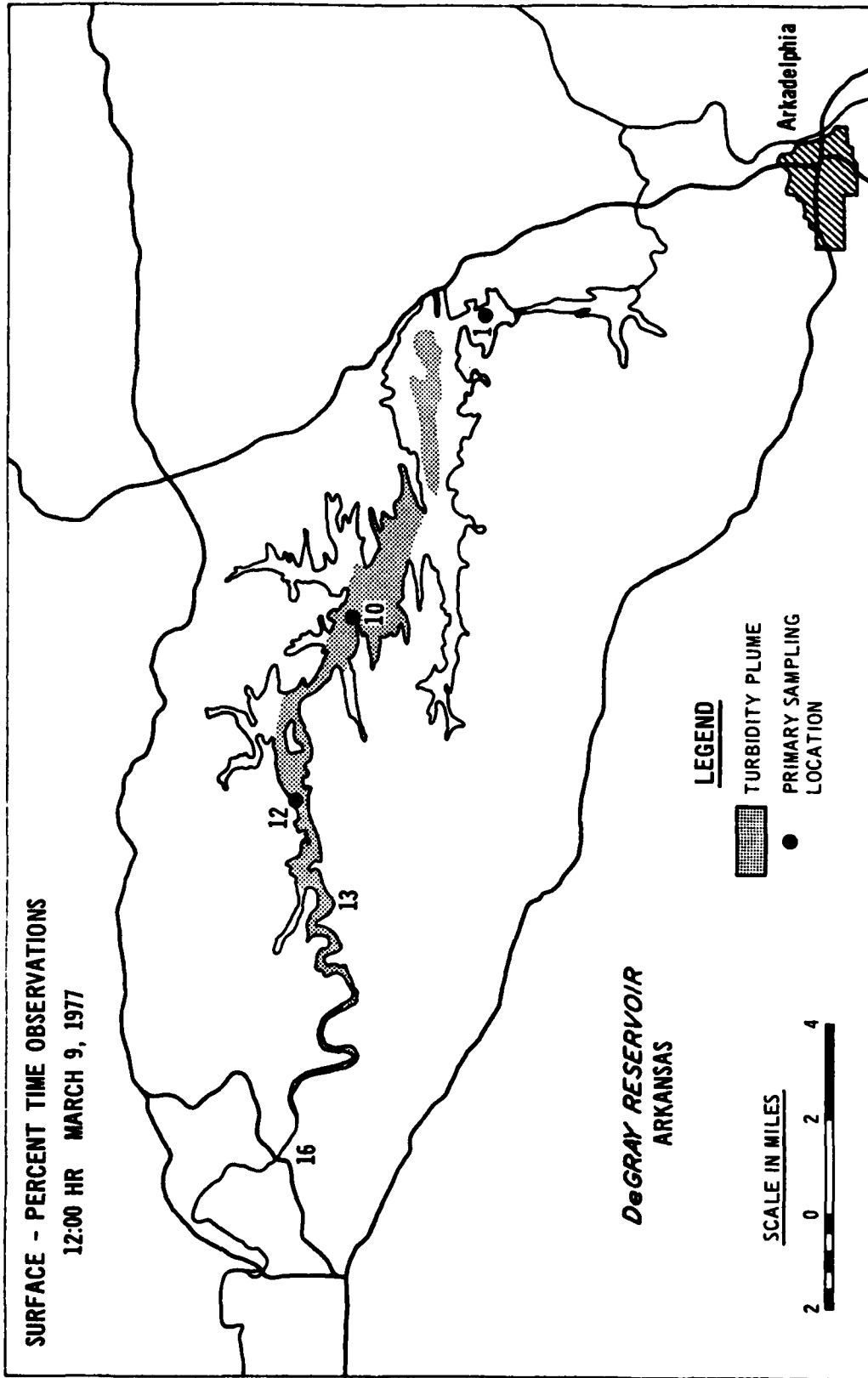


FIGURE 3d

STORM EVENT BEGINNING JUNE 16, 1977

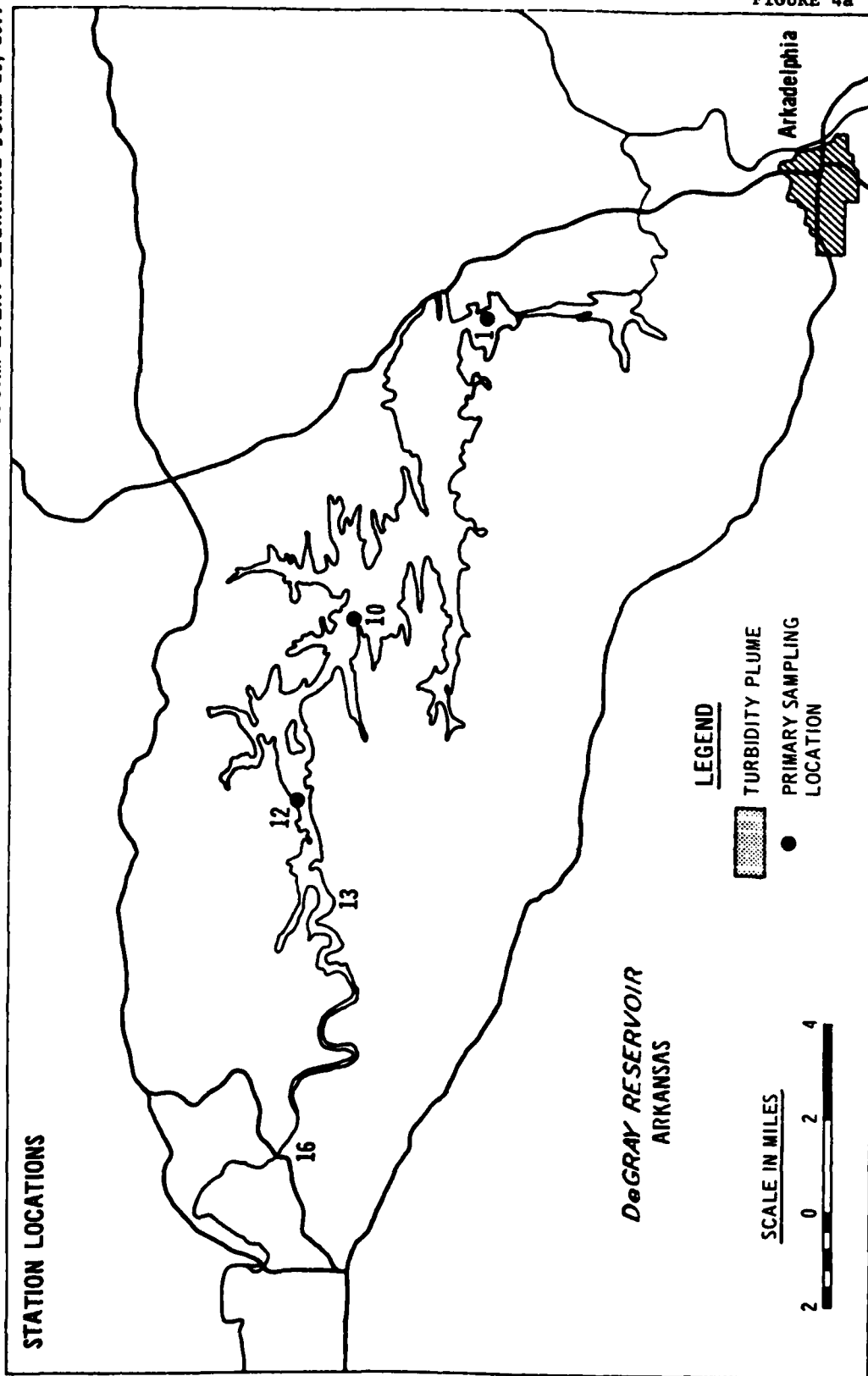
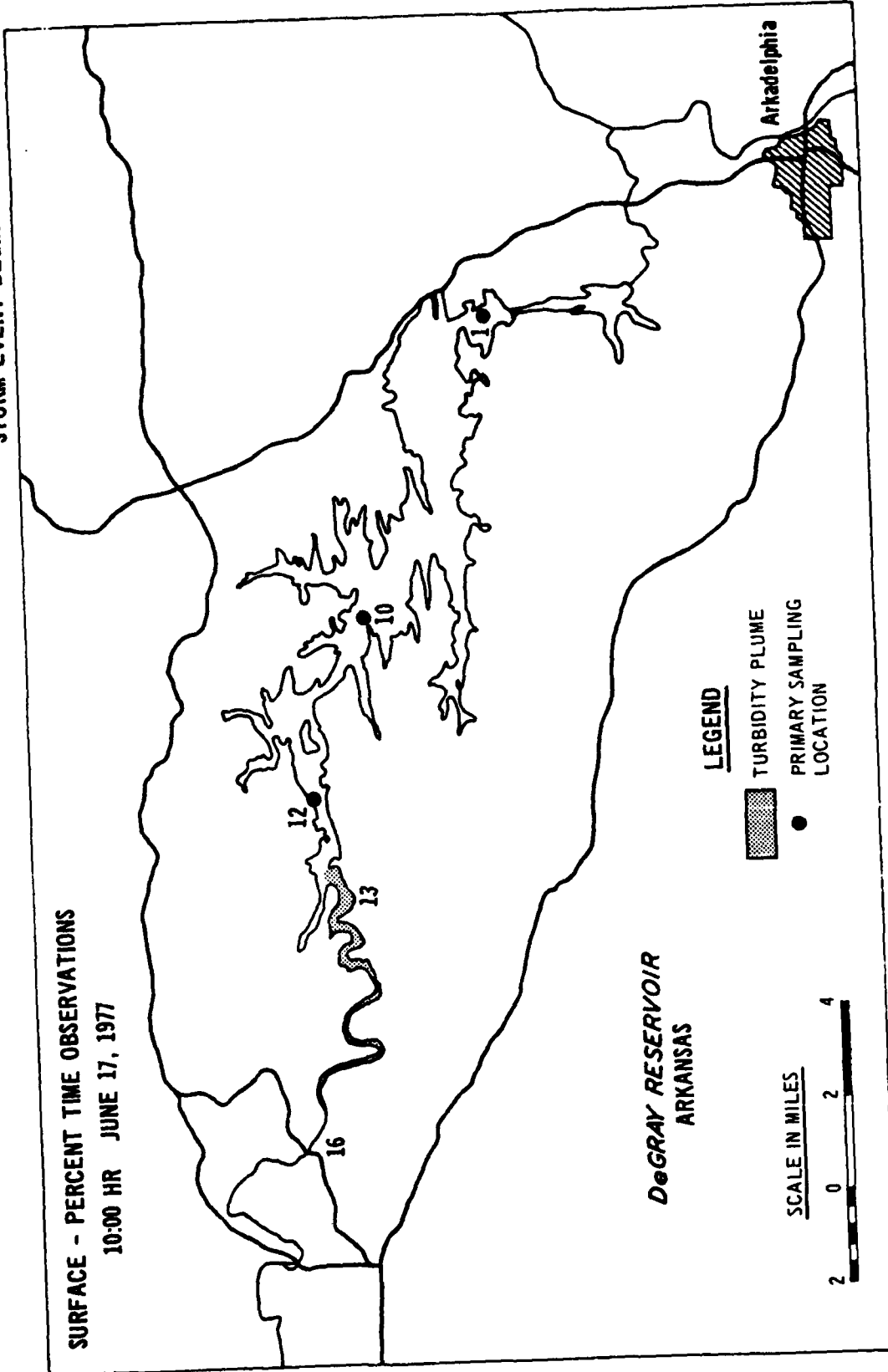


FIGURE 4a

FIGURE 4b

STORM EVENT BEGINNING JUNE 16, 1977

SURFACE - PERCENT TIME OBSERVATIONS
10:00 HR JUNE 17, 1977



STORM EVENT BEGINNING JUNE 16, 1977

SURFACE - PERCENT TIME OBSERVATIONS
12:45 HR JUNE 18, 1977

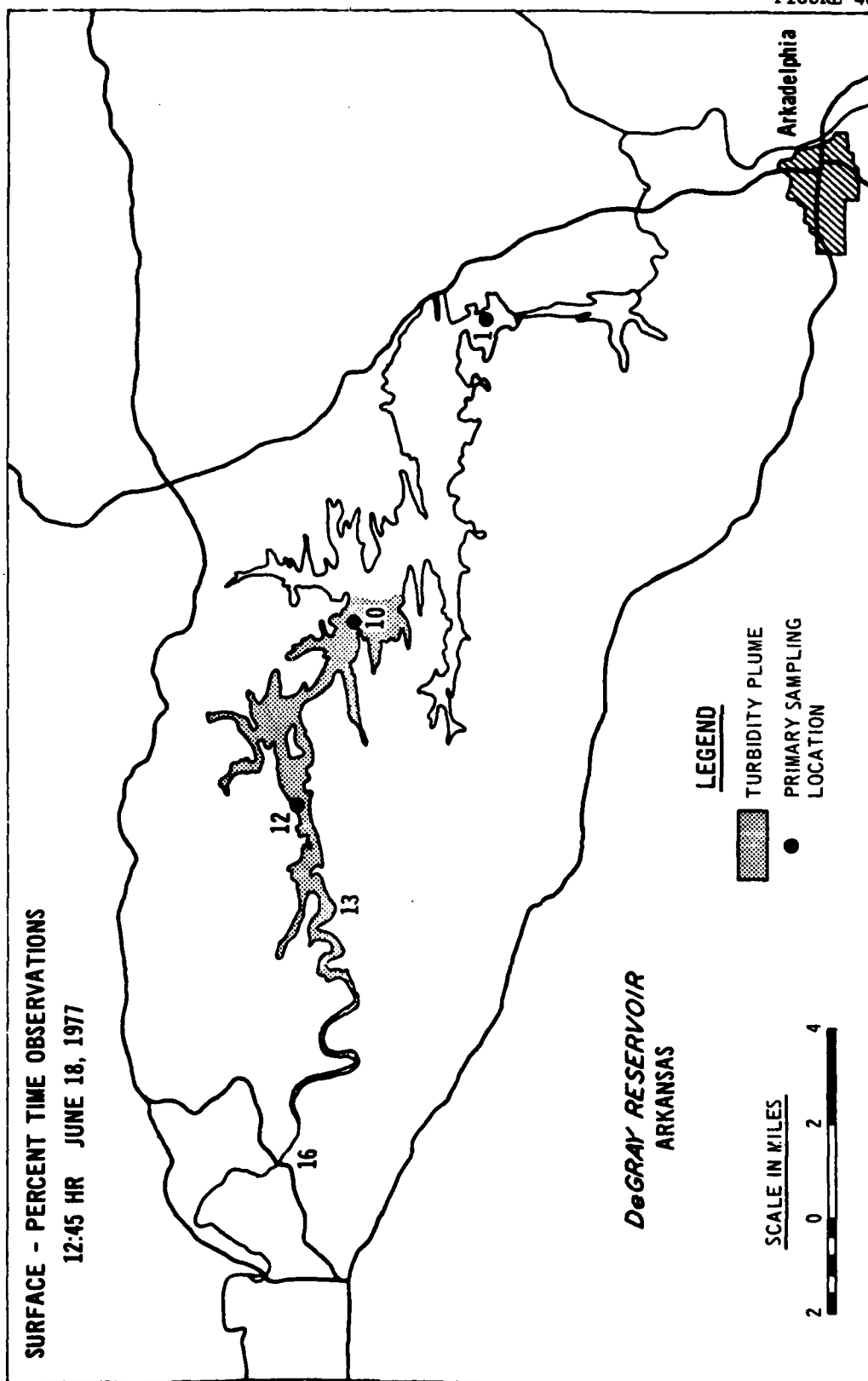


FIGURE 4c

FIGURE 4d

STORM EVENT BEGINNING JUNE 16, 1977

SURFACE - PERCENT TIME OBSERVATIONS
15:35 HR JUNE 20, 1977

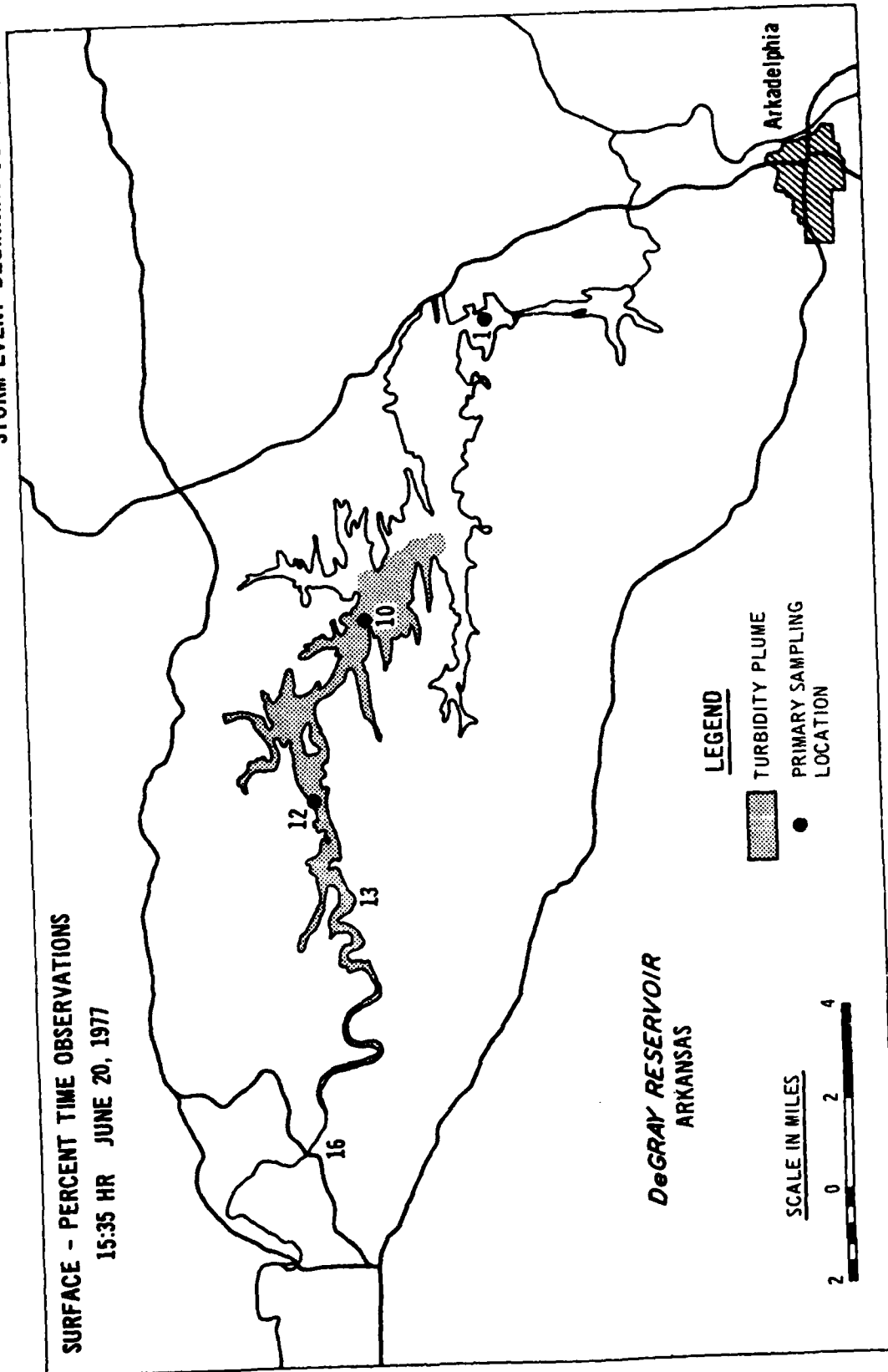


FIGURE 4e

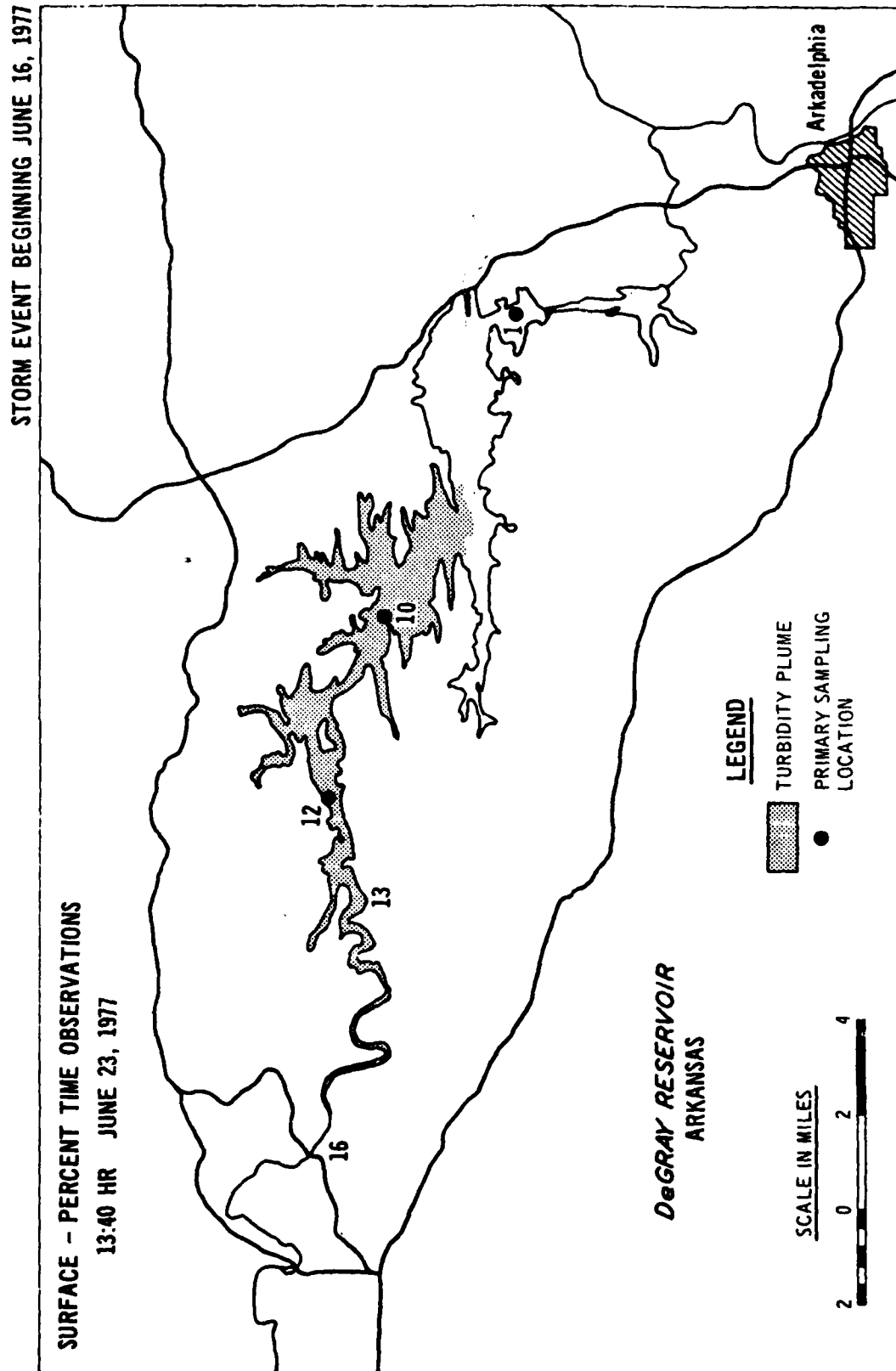


FIGURE 5

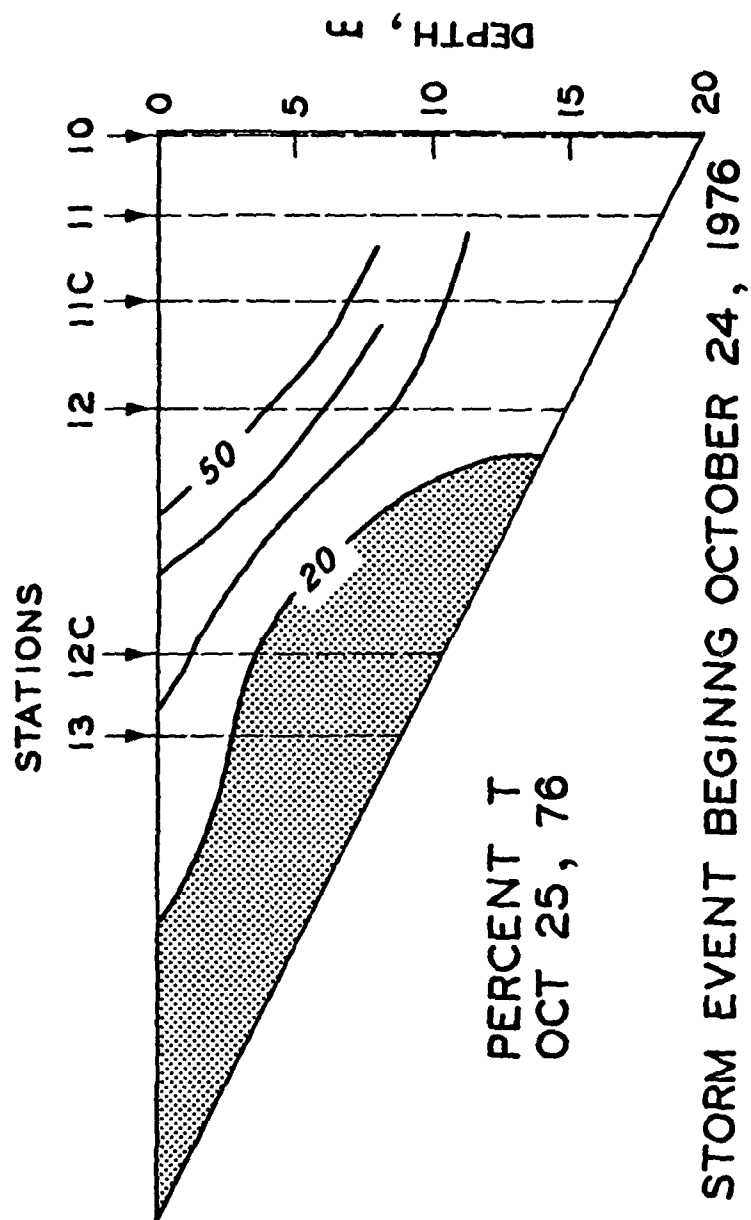


FIGURE 6

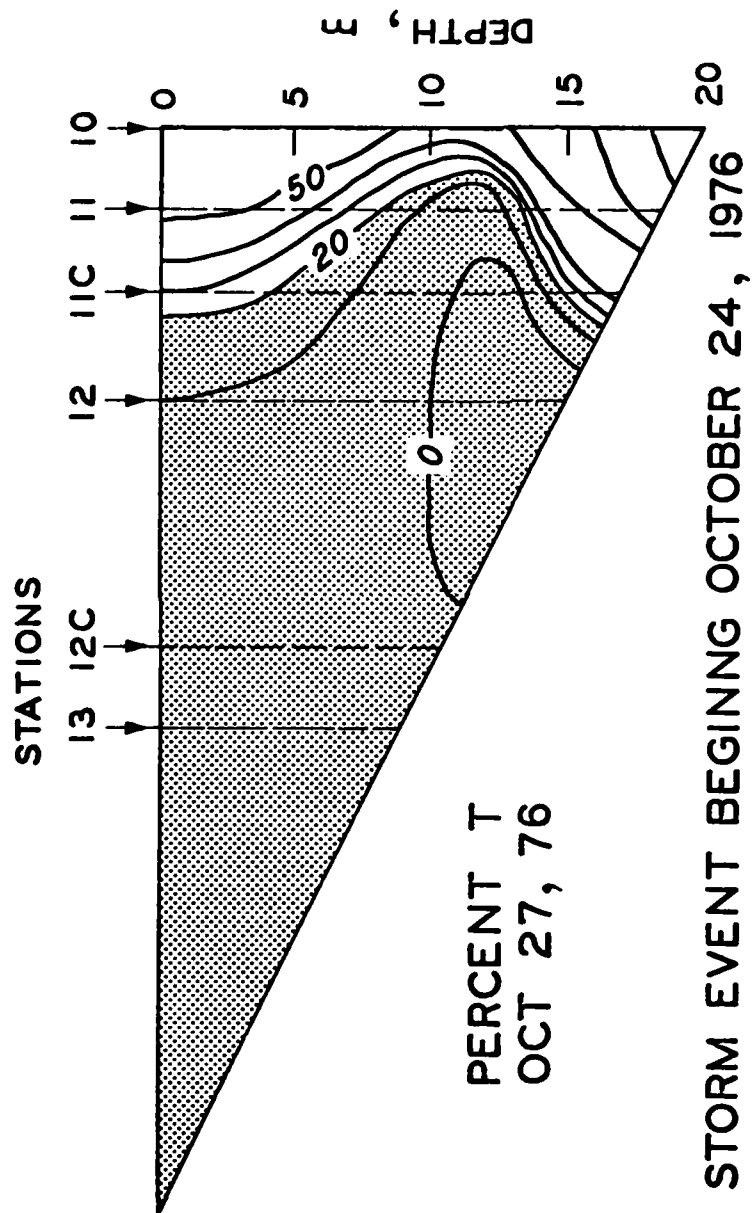


FIGURE 7

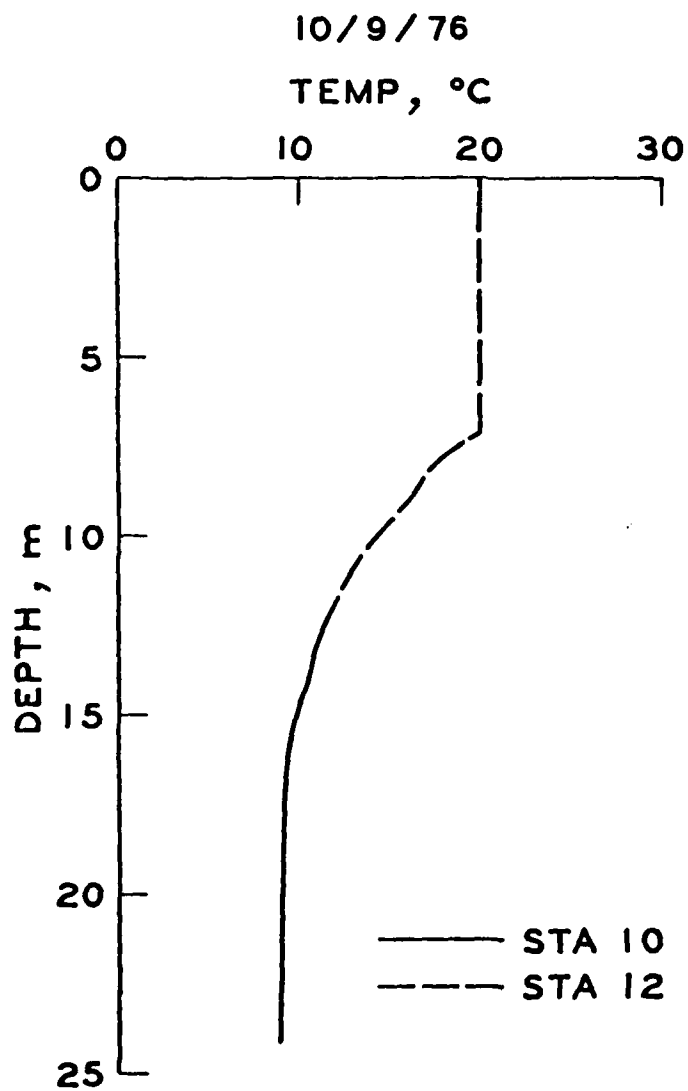


FIGURE 8

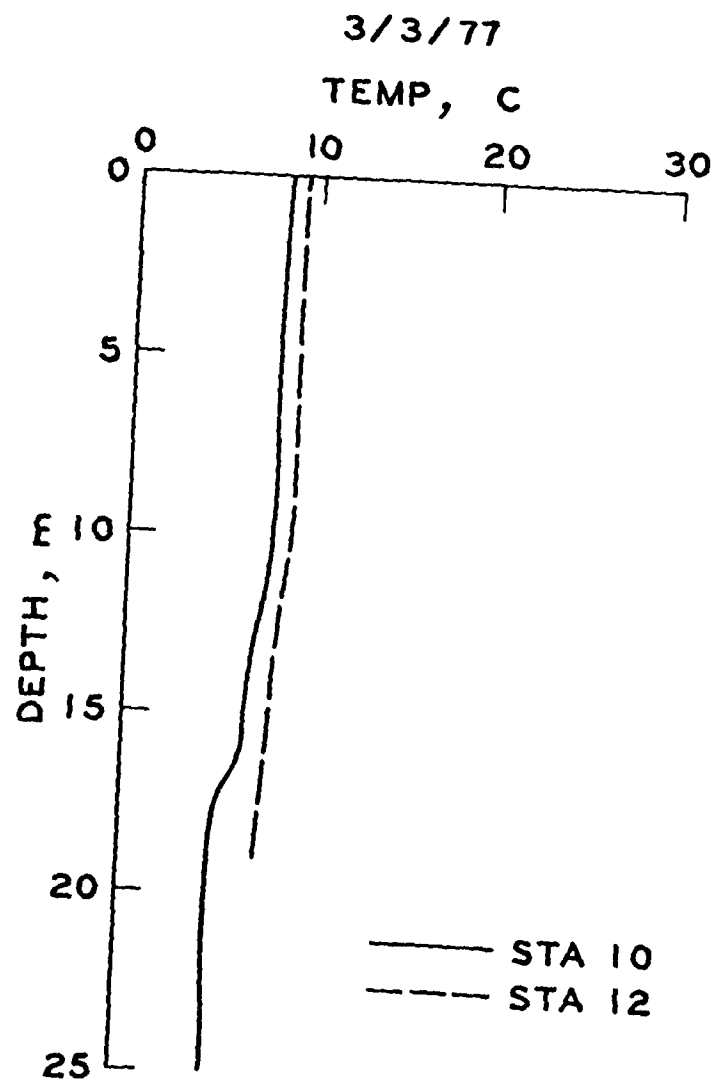


FIGURE 9a

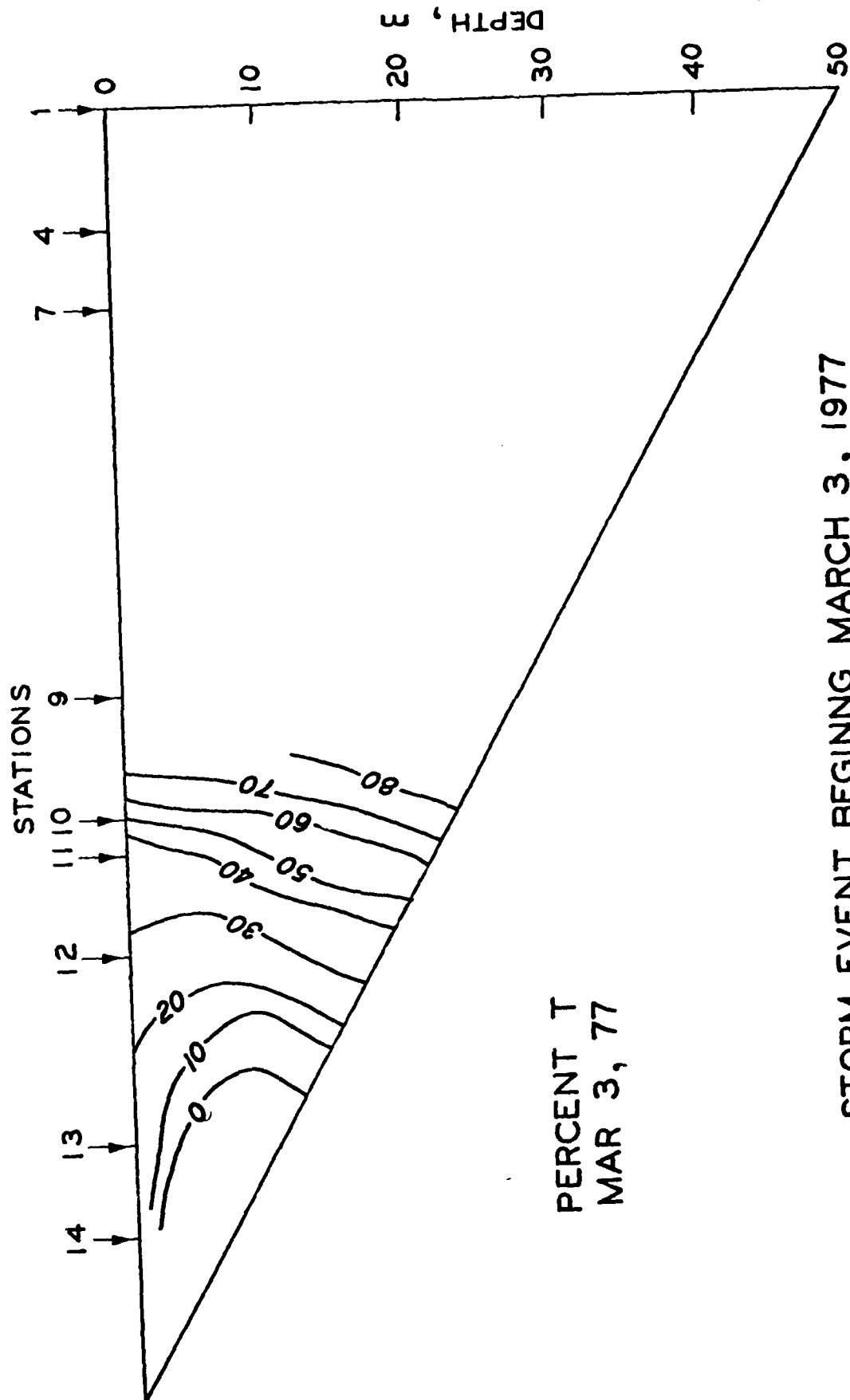
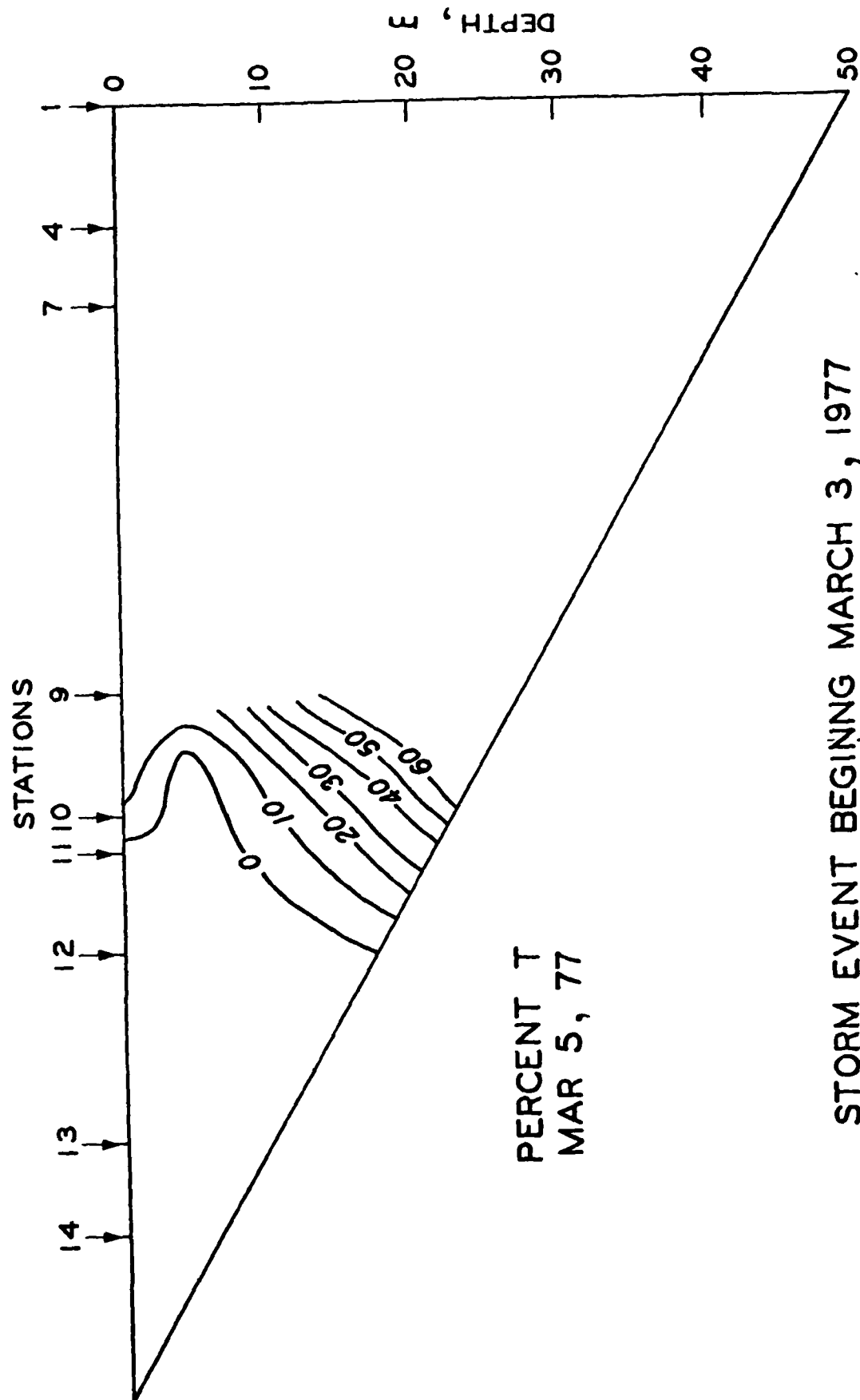
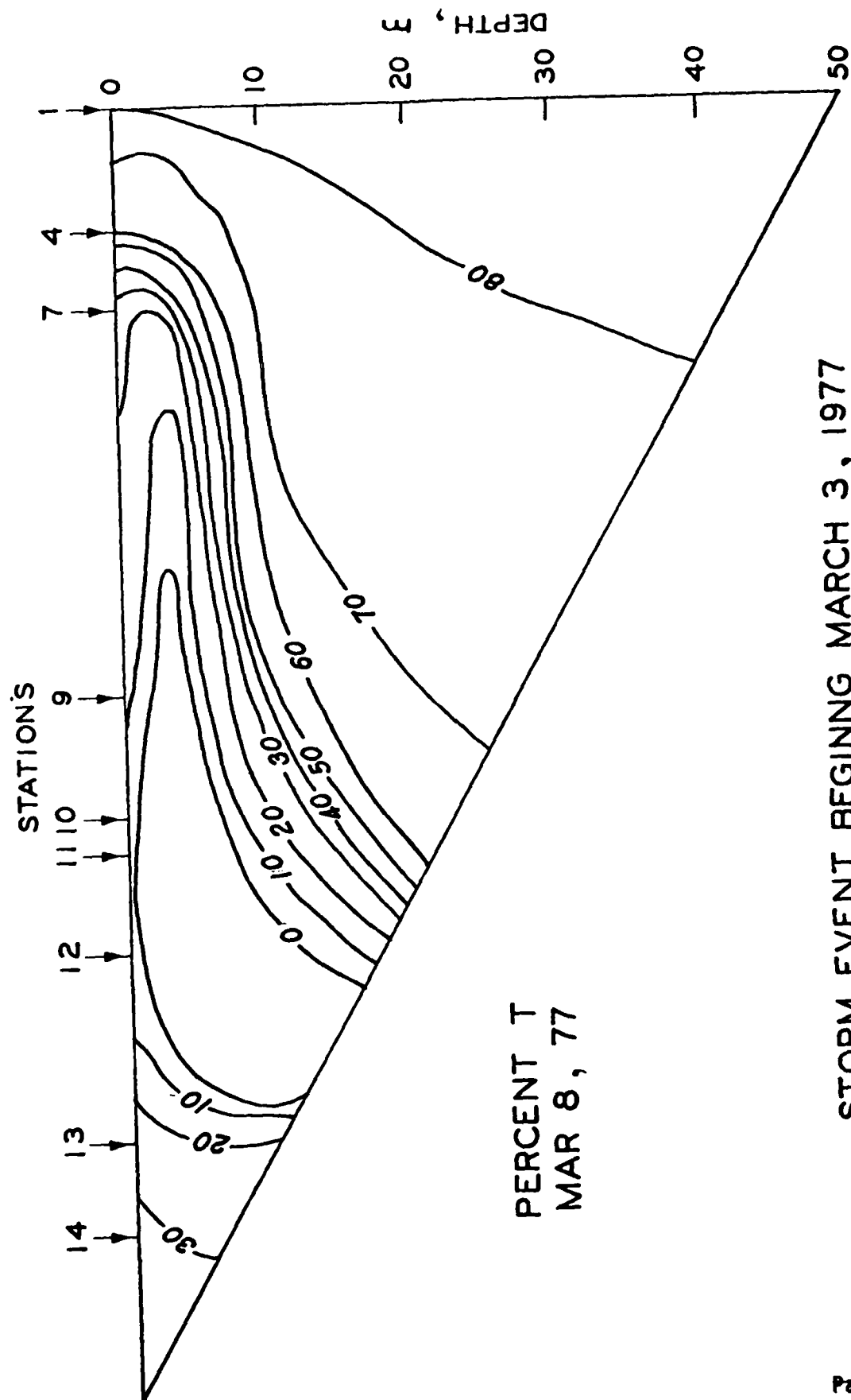


FIGURE 9b



STORM EVENT BEGINNING MARCH 3, 1977

FIGURE 9c



STORM EVENT BEGINNING MARCH 3, 1977

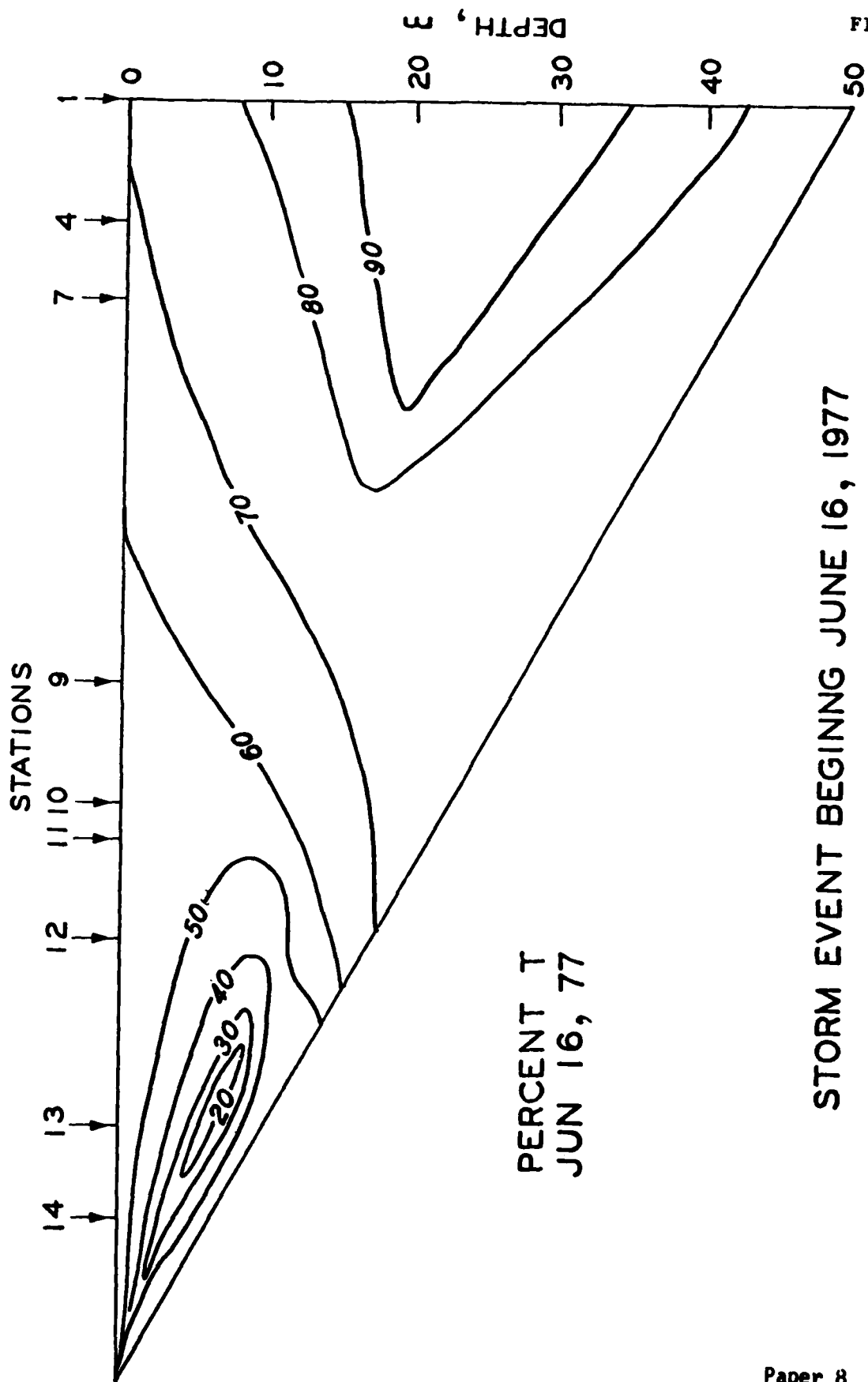
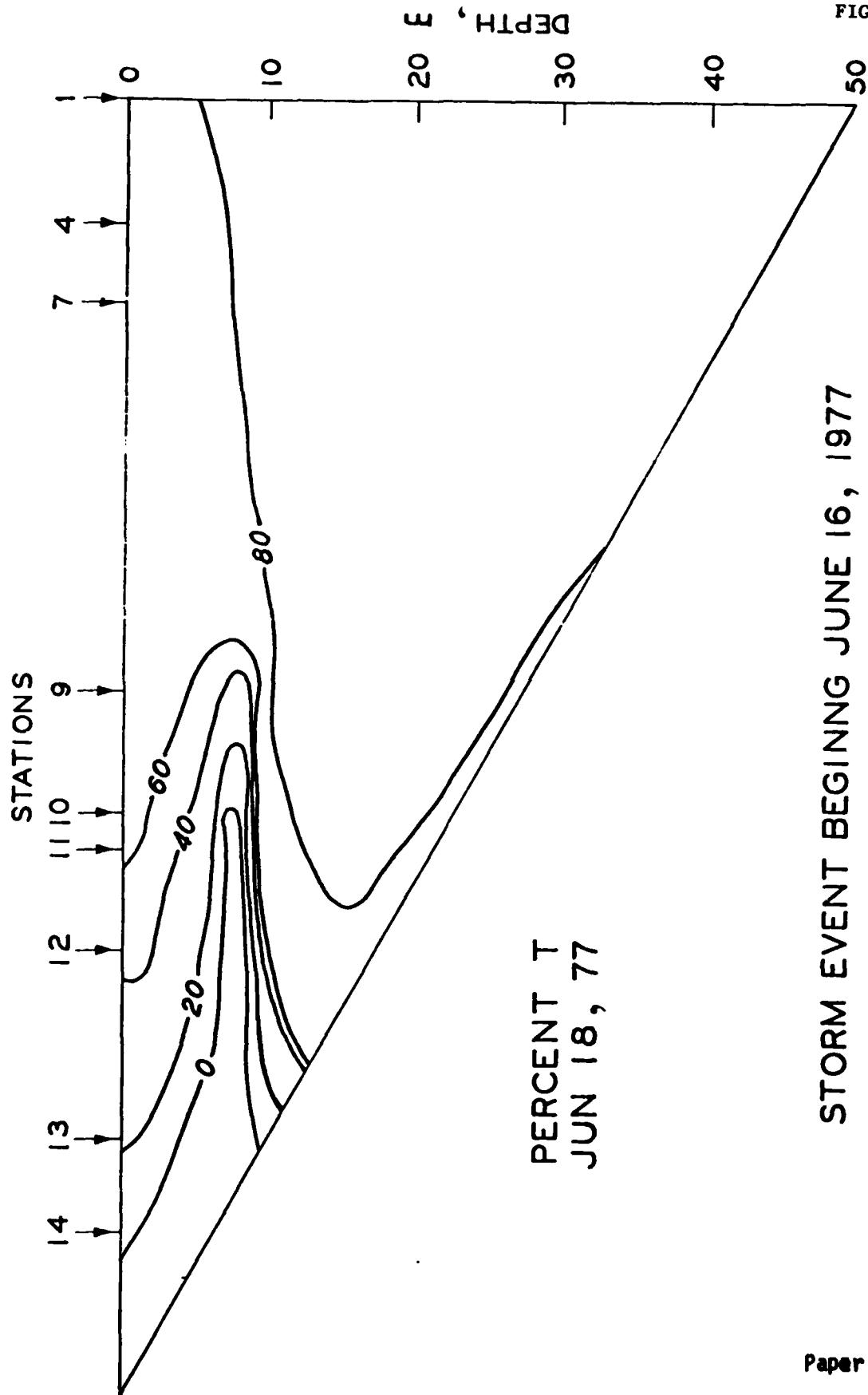
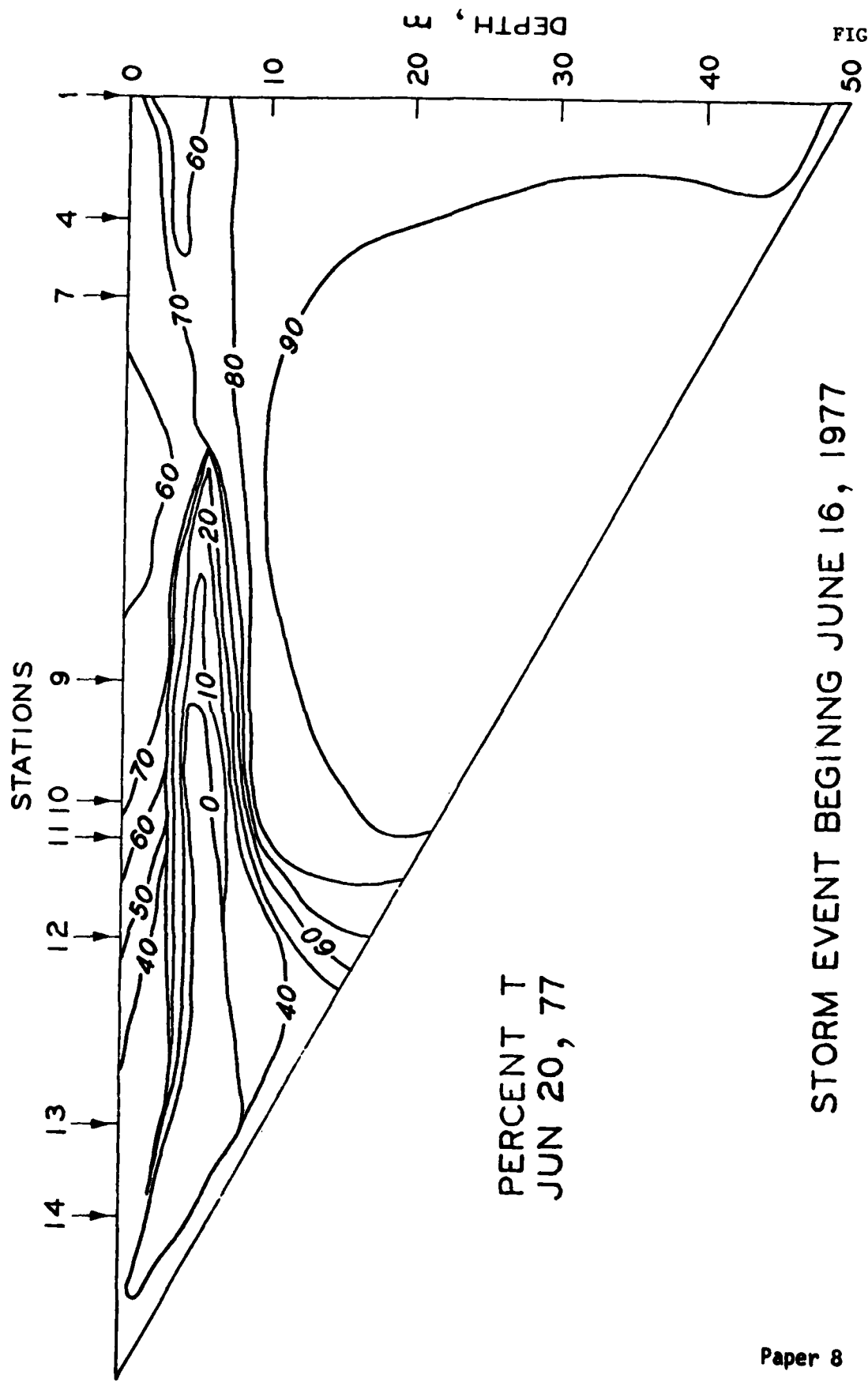


FIGURE 10a

FIGURE 10b

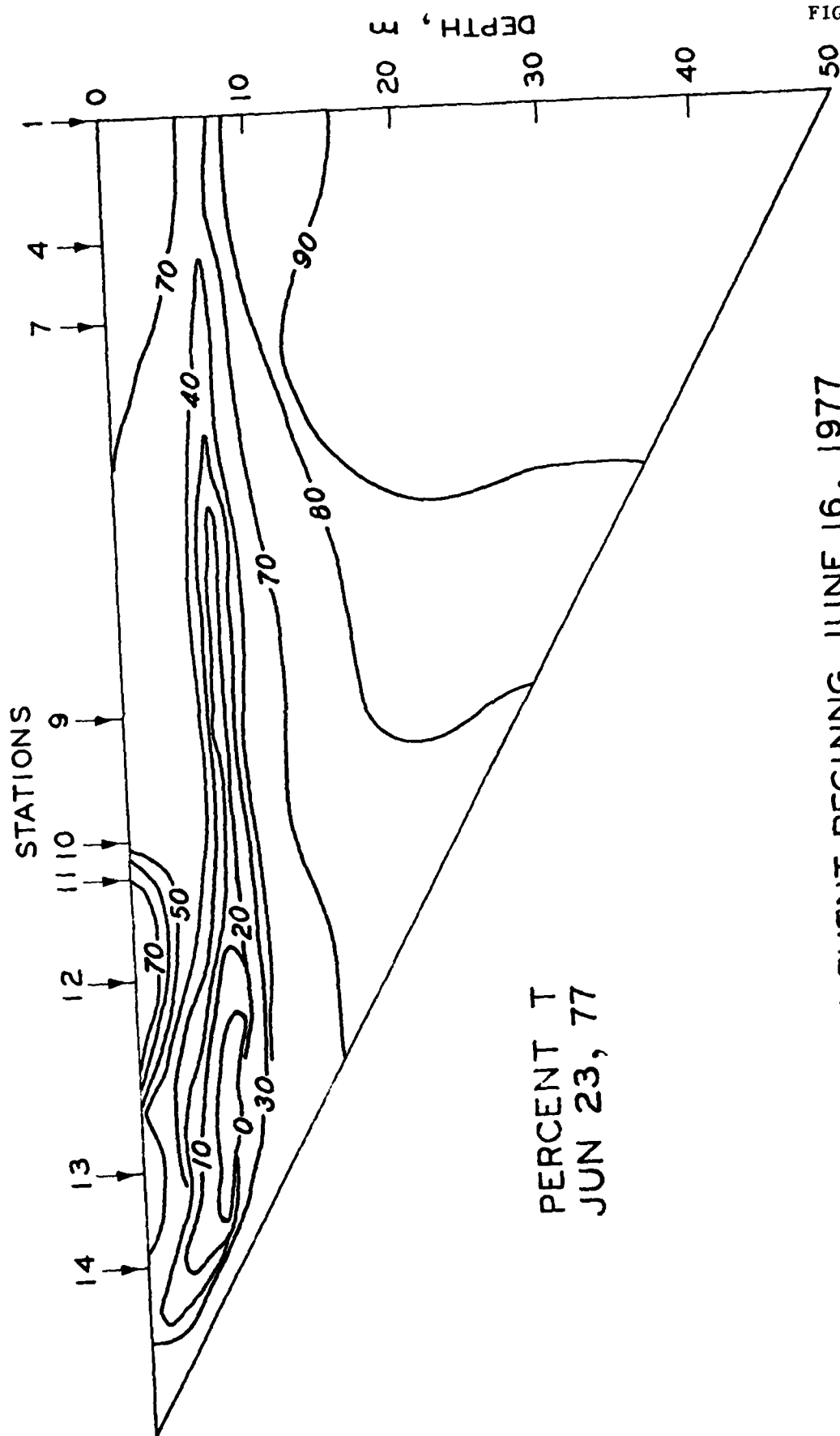




PERCENT T
JUN 20, 77

STORM EVENT BEGINNING JUNE 16, 1977

FIGURE 10c

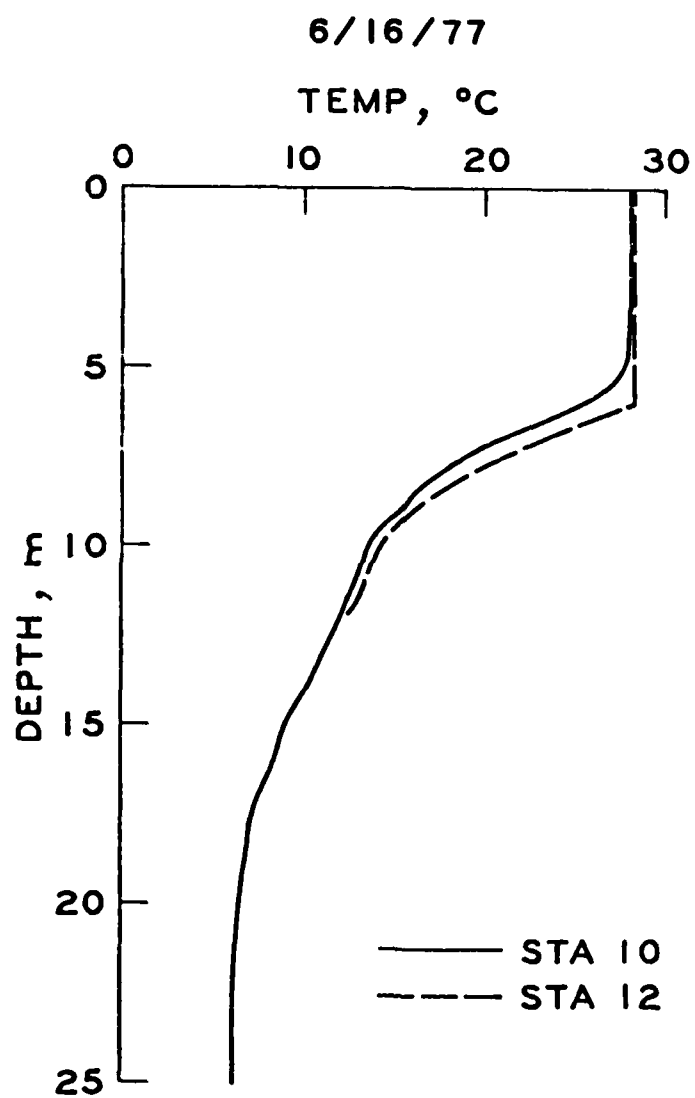


PERCENT T
JUN 23, 77

FIGURE 10d

STORM EVENT BEGINNING JUNE 16, 1977

FIGURE 11



FECAL COLI NO./100 ML, $\times 10^3$

1973

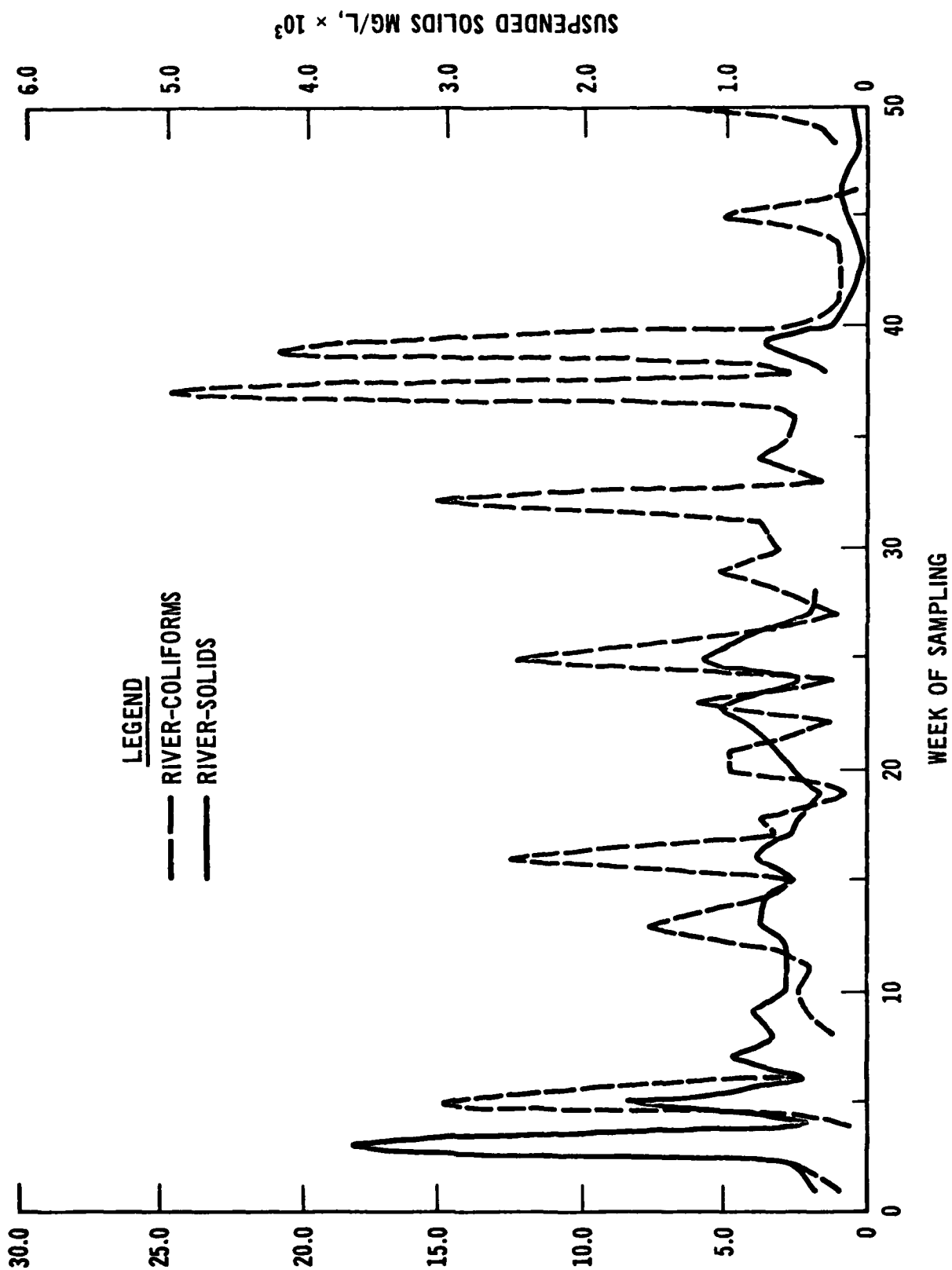


FIGURE 12

FIGURE 13a

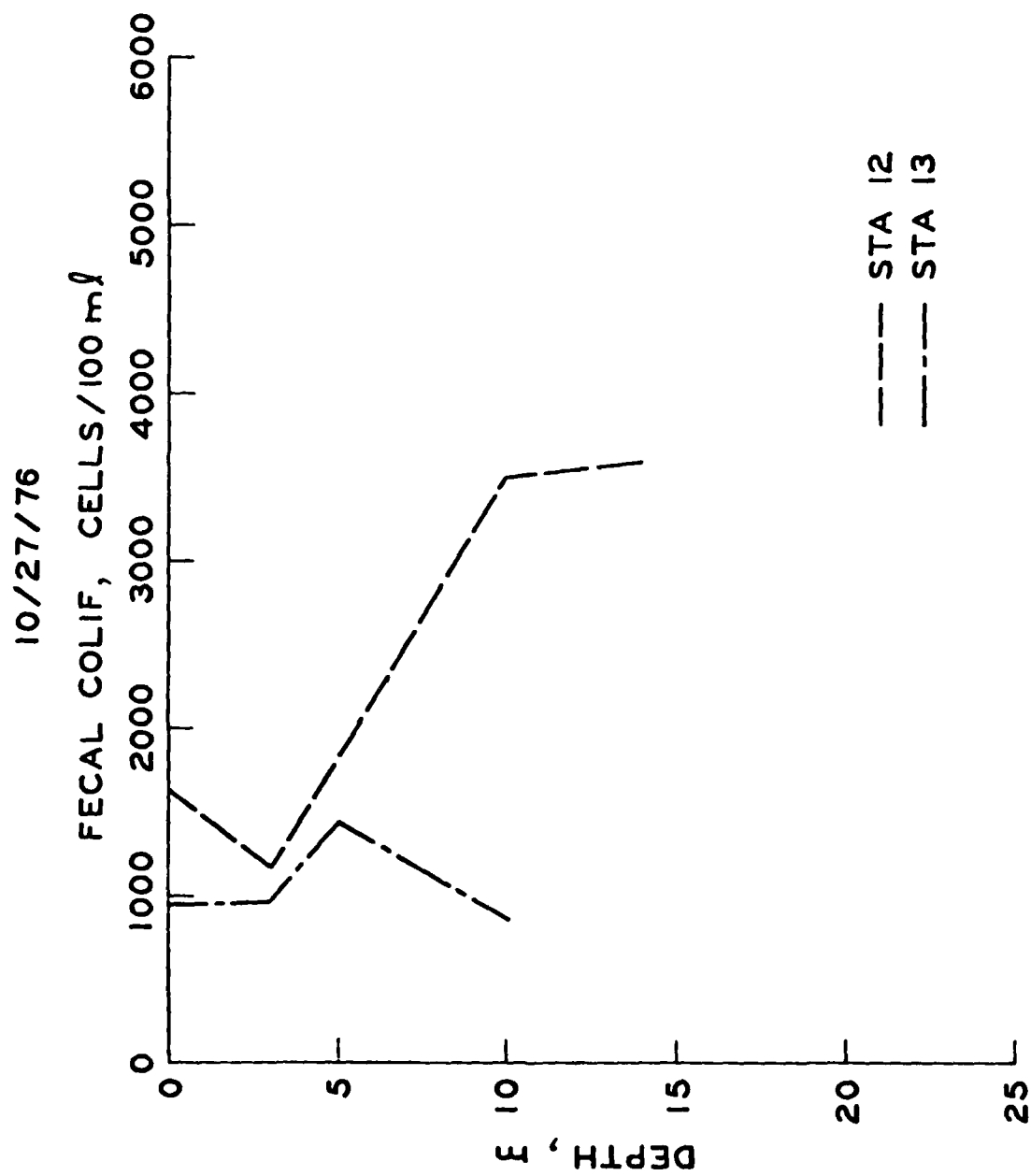


FIGURE 13b

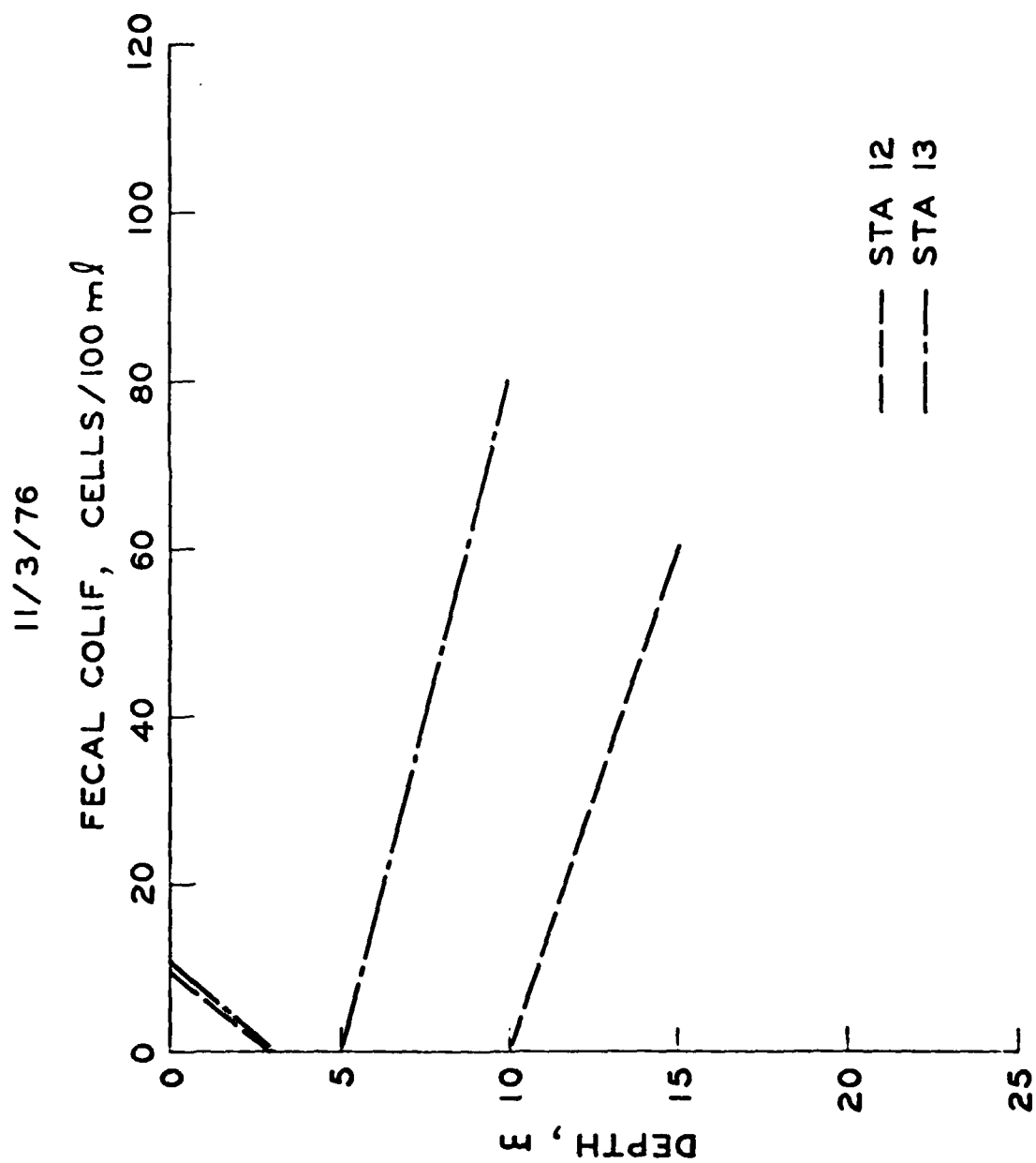


FIGURE 14a

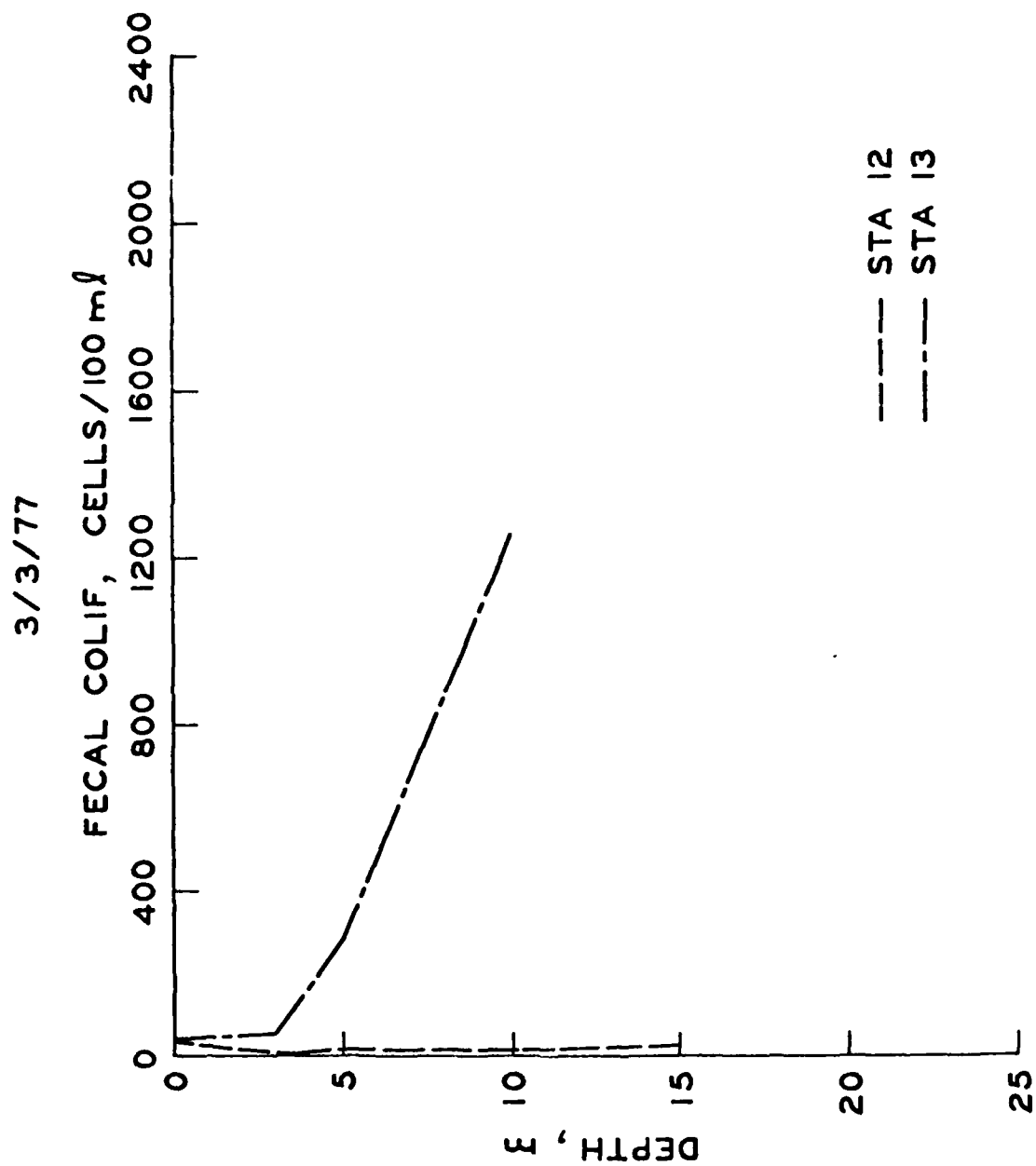


FIGURE 14b

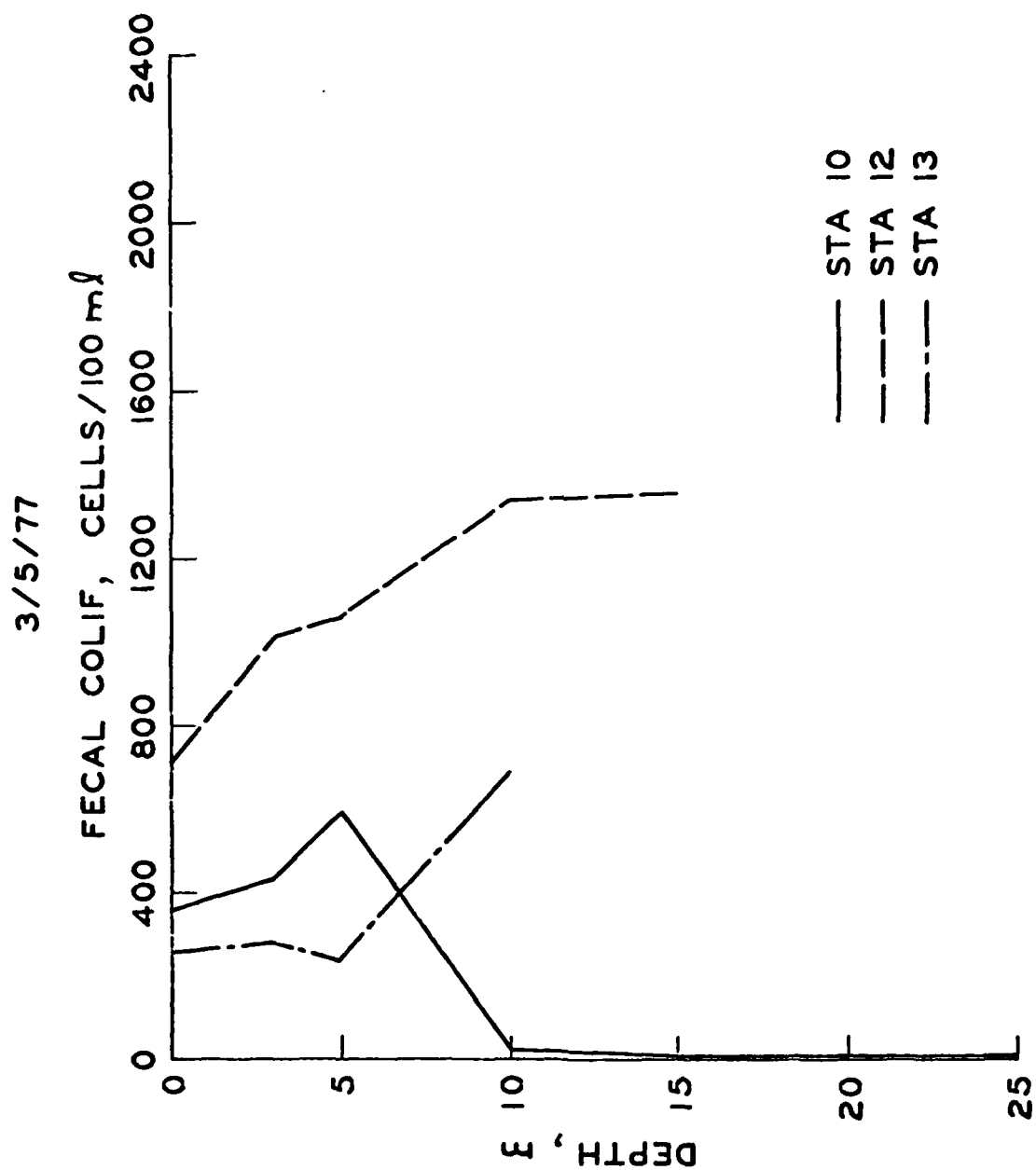


FIGURE 15a

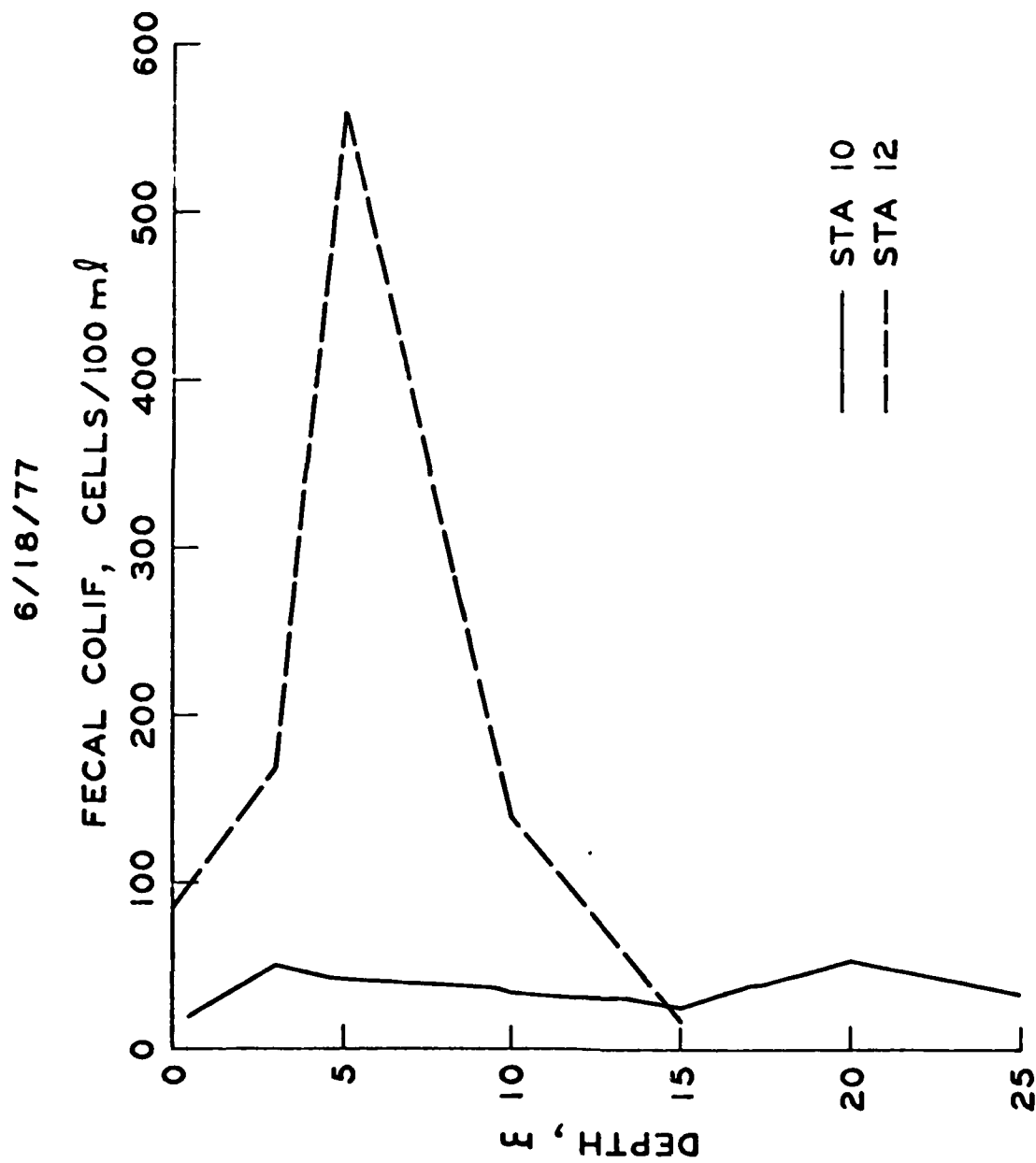


FIGURE 15b

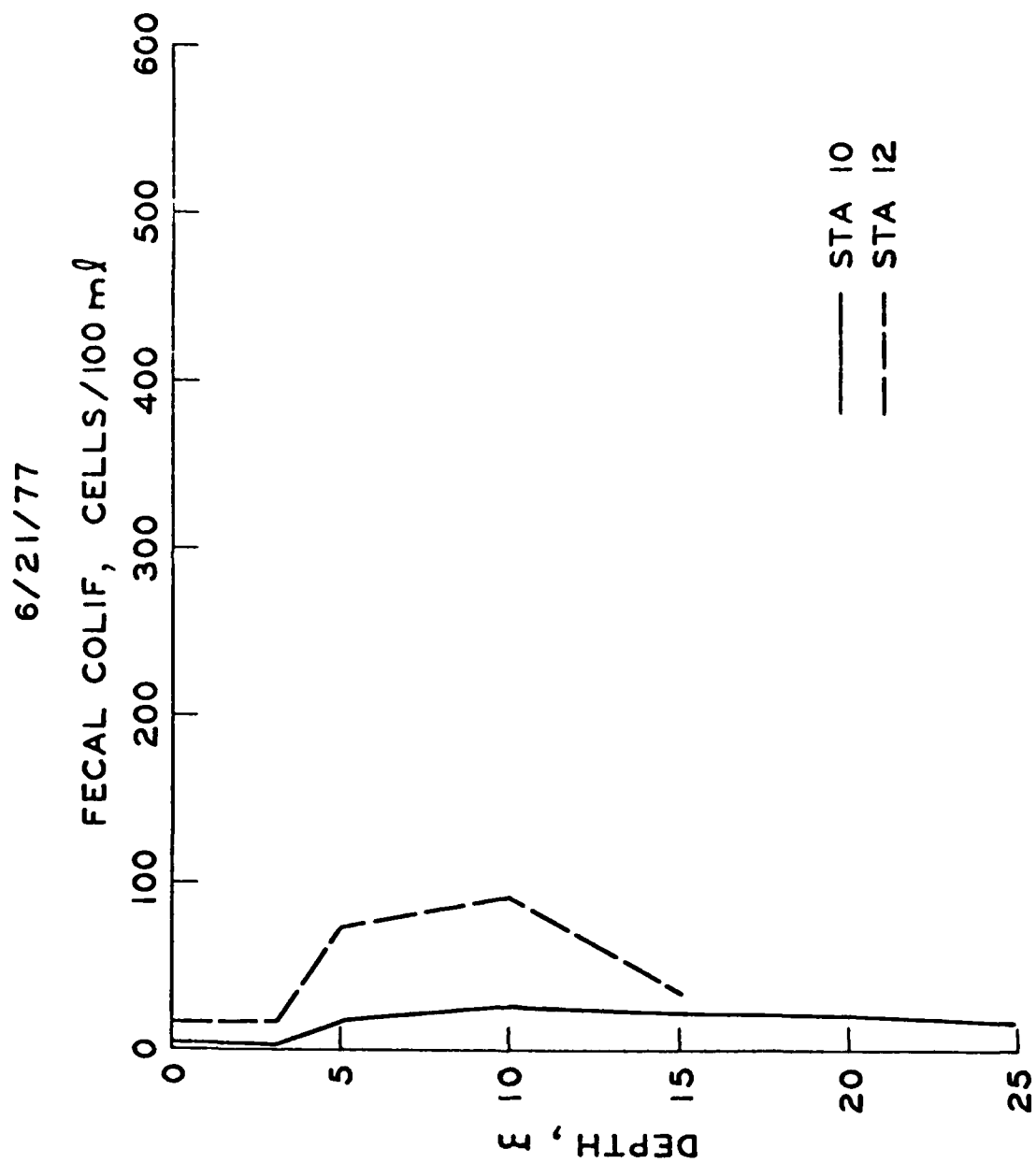


FIGURE 16

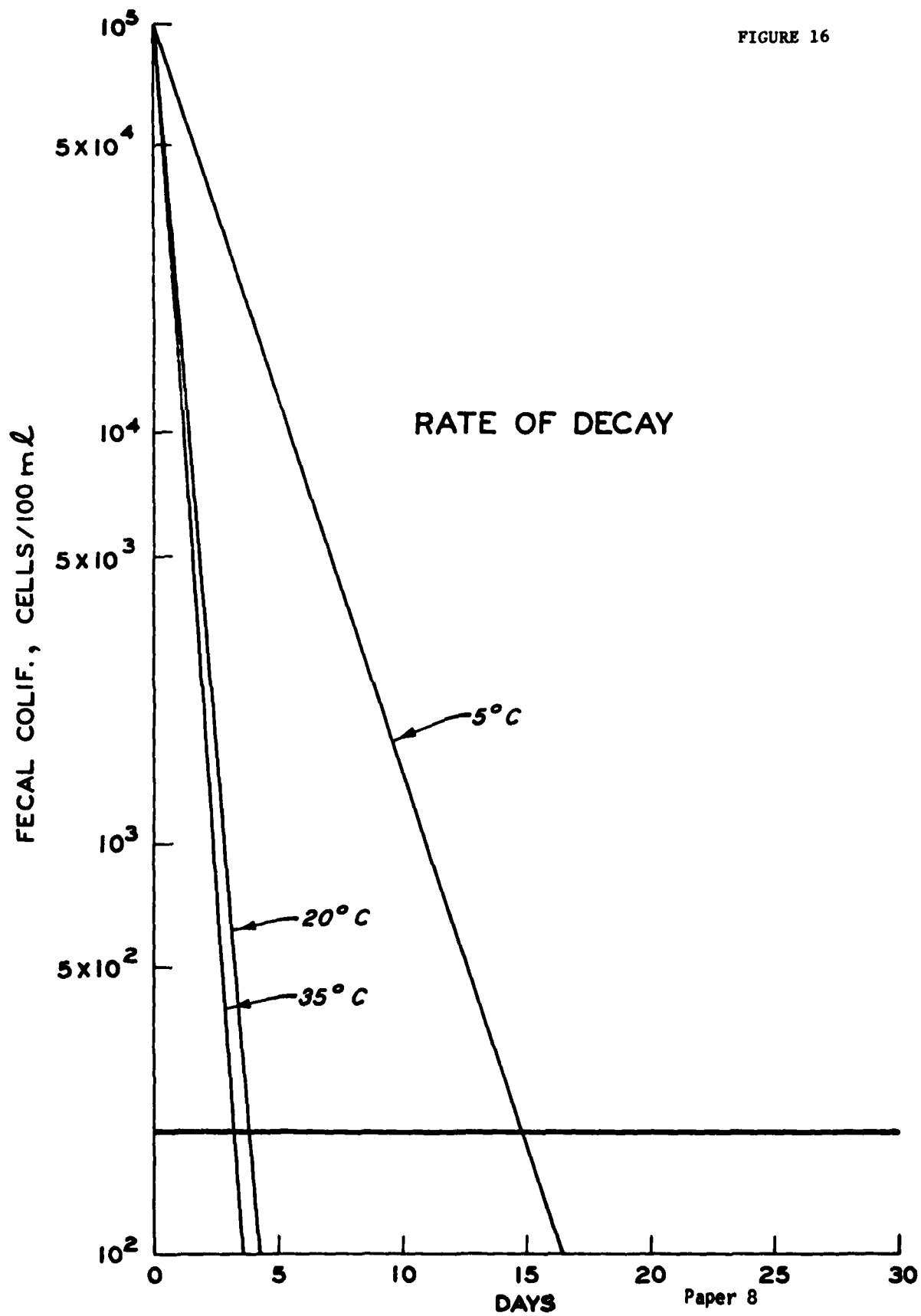
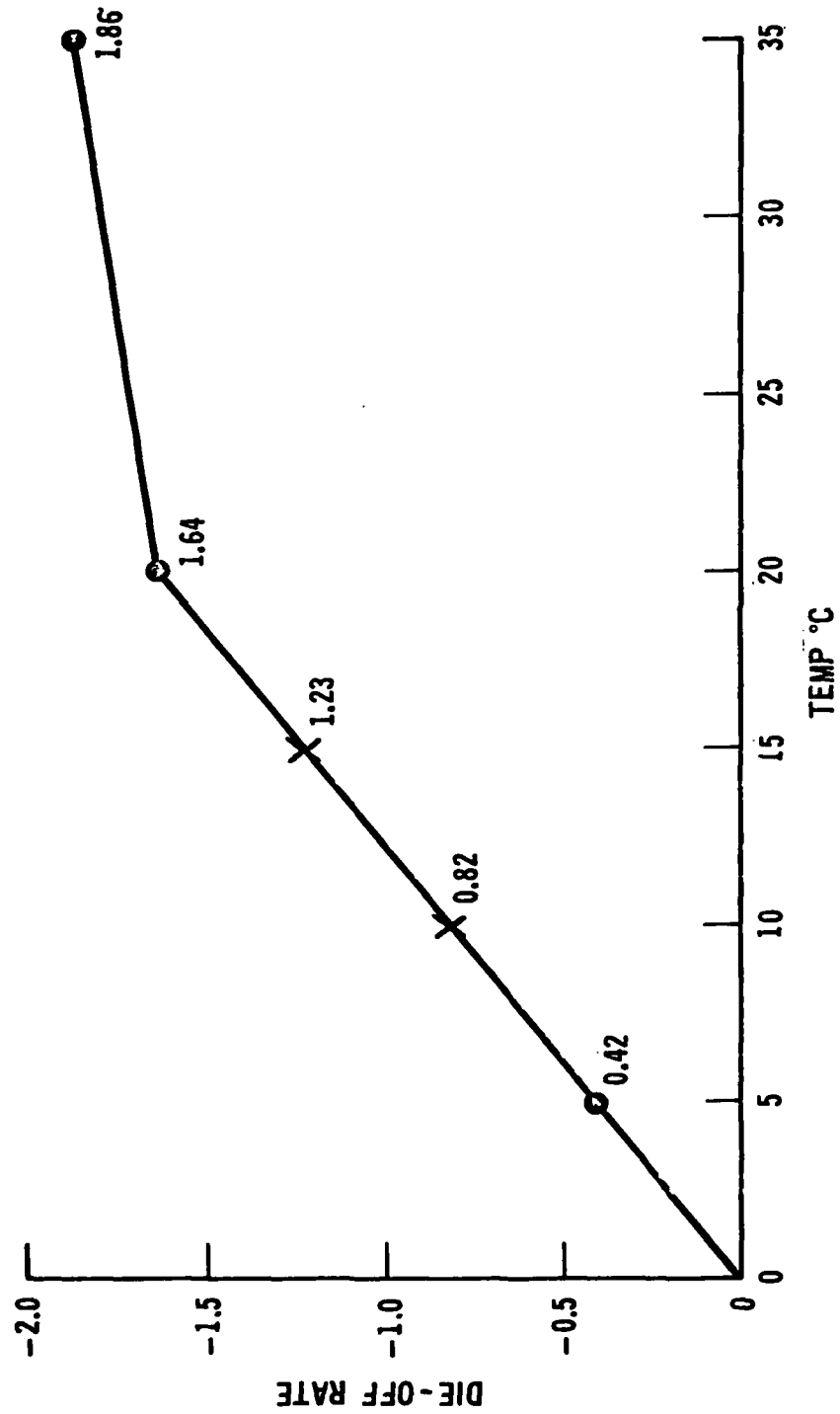


FIGURE 17



AQUATIC BIOLOGY - DATA INTERPRETATION
(PART I): PRIMARY PRODUCER ORGANISMS
(ALGAE AND MACROPHYTES)

By

John W. Barko, Research Biologist
Environmental Effects Laboratory
Waterways Experiment Station
Vicksburg, MS U.S.A.

INTRODUCTION

The process of photosynthesis utilizes solar energy to convert carbon into organic matter, and thereby satisfies the nutritional and energetic requirements of most organisms within the biosphere. Organisms capable of photosynthesis are referred to as primary producers. These organisms are autotrophic - meaning that they require only a source of kinetic energy (light) and certain essential nutrients for self-maintenance. Through their metabolism, primary producers considerably influence the quality of the environment.

In aquatic ecosystems, the primary producers are the algae and macrophytes ("large plants"). Because of their significant interactions with other components of the aquatic ecosystem, these organisms markedly affect water quality. Therefore, an understanding of the role of aquatic primary producers is appropriate for the proper evaluation and interpretation of water quality data.

Interactions among aquatic primary producer organisms, other aquatic organisms, and the physical and chemical aquatic environment are discussed herein. This article provides basic information on these interactions and elucidates some of the limitations and constraints associated with related methods of parameterization.

ECOLOGICAL CLASSIFICATION
OF AQUATIC PRIMARY PRODUCERS

Aquatic macrophytes are simply large aquatic plants. Although the term "macrophyte" is most commonly used in reference to the more structurally complex aquatic flowering plants, certain large algae

and also aquatic bryophytes (mosses) are also referred to as macrophytes. For example, the macroalga, Chara, and the moss, Fontinalis are macrophytes. Aquatic macrophytes can be classified into four categories on the basis of their life form.

| Life Form | Example |
|-----------------|------------|
| emergent | cattail |
| floating-leaved | water lily |
| submergent | pondweed |
| free-floating | duckweed |

Members of the first three categories (emergent, floating-leaved, and submergent plants) are rooted and attached to a substratum. Free-floating plants often have roots, but these do not function in attachment. Emergent aquatic plants, although adapted to constant or periodic standing water, are most similar to terrestrial plants. These have easily distinguishable and functionally separable aerial and rooted portions. Submergent plants are the most highly adapted to the aquatic environment. These plants remain totally submersed throughout their lifespan, lack structural rigidity, and therefore are dependent upon the buoyancy of the water for their support. Floating-leaved plants are intermediate in form between emergent and submergent life forms, and have characteristics of both. Similar to the emergent aquatic plants, floating-leaved plants possess separate aerial and rooted portions. In addition to the large floating leaves, smaller leaves similar to those of submerged plants are often produced at the sediment surface.

Algae are structurally simplified aquatic plants. In general, they lack functionally differentiated structures (i.e., roots, leaves, etc.) and are relatively (compared to macrophytes) small in size. This criterion of size necessitates the classification of macroalgae as macrophytes. Because of the obvious ambiguities involved, the term microphyte, meaning small aquatic plant, is occasionally substituted for the term, algae.

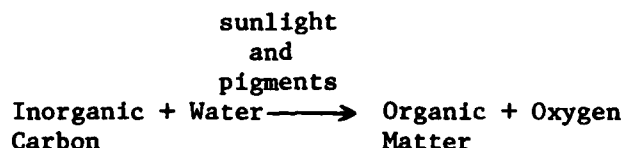
Algae can be ecologically classified on the basis of their mode of occurrence. In this respect, there are two major categories: phytoplankton and attached algae. Phytoplankton are free-floating, and inhabit the pelagic region (i.e., open water) of bodies of water. Attached algae, often called periphyton, are physically associated with a substratum. Their attachment can be simply mechanical, or it may involve a form of mutualism. For example, there is some evidence indicating that epiphytic algae (i.e., algae attached to other plants)

may derive some of their nutrition from the host plant. In contrast, the attachment of algae to rock surfaces is usually simply mechanical. Attached algae are easily dislodged by water movements; thus, they are frequently found in the phytoplankton, but normally so in relatively low abundance. Also, some algal species exist abundantly in both the planktonic and the attached modes of occurrence. Attached algae are frequently further classified according to the type of substratum with which they are associated. Commonly used examples of this type of classification are given below.

| <u>Type of Attachment</u> | <u>Substratum</u> |
|---------------------------|-------------------|
| epipellic | fine sediments |
| epilithic | rocks |
| episammic | sand grains |
| epiphytic | plants |
| epizoic | animals |

INTERACTIONS BETWEEN AQUATIC PRIMARY PRODUCERS AND THE PHYSICAL ENVIRONMENT

Sunlight is used by primary producers in the process of photosynthesis, presented below in a simplified manner.



In photosynthesis, sunlight is absorbed by various pigments. Among the many pigments used in this process, chlorophyll a is the most important because it is directly involved in the conversion of solar energy to chemical energy as organic matter is synthesized from inorganic carbon. Only chlorophyll a is found in all primary producers. Other pigments include carotenoids, phycobolins, and other chlorophylls. These other pigments are termed "accessory pigments" because they transfer absorbed light energy to chlorophyll a. On a dry weight basis, the mass of chlorophyll a in plant cells is usually much greater than the mass of other pigments. Changes in the relative proportions of different pigments (i.e., chromatic adaptation) occur in response to changes

in the quality of light as explained below. Thus, both the color and light absorption capability of a single species can differ among different habitats.

Light of wavelengths between approximately 390 and 760 nanometers is used in photosynthesis. This range represents the violet to red portion of the visible light spectrum. In general, the penetration of different wavelengths of light into natural water depends upon the quantity and kinds of dissolved materials and suspended particulate materials that differentially absorb and reflect portions of the light spectrum. In many reservoirs, light penetration is significantly limited by turbidity associated with suspended sediment. The same effect can occur during phytoplankton blooms that may be dense enough to also cause turbidity. Turbidity limits primary productivity by decreasing light penetration and consequently limiting photosynthesis.

Water temperature affects the rates of various biochemical processes. In general, overall cellular metabolism is temperature-dependent. Different macrophyte and algal species exhibit different temperature optima. Temperature and light interact to define four hypothetical categories of primary producer organisms, each adapted to a specific season (see below).

| | low light | high light |
|------------|-----------|------------|
| low temp. | Winter | Spring |
| high temp. | Autumn | Summer |

Ignoring for the moment other important factors (i.e., chemical, biological, and other), temperature and light collectively account for the seasonality of occurrence of primary producer organisms.

Indirect effects of temperature on primary producers are complex, but nonetheless important. Changes in the density of water with water temperature change affect phytoplankton buoyancy, and ultimately the vertical distribution of phytoplankton. Temperature also affects the availability of nutrients. Rates of microbial mineralization of organic matter and subsequent releases of nutrients are greater at higher than at lower temperatures. However, the solubility of gases including carbon dioxide, an important substrate in photosynthesis, is inversely related to water temperature.

The majority of macrophytes are rooted, and their growth and distribution are markedly affected by sediment type. The role of sediments in the nutrition of macrophytes is variable, because nutrients may be absorbed either from the water by the shoots or from the sediment by the roots. For the most part, regions of nutrient absorption

depend upon the relative availability of nutrients with respect to the distribution of plant tissues having an absorptive capability. Nutrient uptake by emergent macrophytes, which have extensive root systems and little absorptive surface exposed to the water, probably occurs almost exclusively from the sediment. In the case of submergent macrophytes, recent evidence has suggested that their roots are functionally similar to those of emergent life forms. Therefore, because of the usually greater availability of nutrients in sediments than in water, sediments represent a potentially important source of nutrition to these macrophytes. Primarily for nutritional reasons, macrophyte communities achieve maximum growth on fine-textured sediments.

By virtue of their ability to colonize fine-grained materials, rooted macrophytes promote sediment stabilization. Dense beds of macrophytes are also capable of reducing turbidity by decreasing the velocity of water currents that pass through them; this consequently increases rates of sedimentation. Macrophytes significantly affect the shoreline regions of hard water lakes by promoting the development of marl (calcium and magnesium carbonate) formations that gradually expand the area of shallow regions at the expense of the open water zone.

INTERACTIONS BETWEEN AQUATIC PRIMARY PRODUCERS AND THE CHEMICAL ENVIRONMENT

Various nutrients are essential to primary producers. Carbon, nitrogen, and phosphorus are among those nutrients that are required in large concentrations, and are referred to collectively as macronutrients. Micronutrients include inorganic elements and organic compounds such as metals and vitamins, respectively, that are required in small concentrations. Nutrients required by aquatic primary producers are basically the same as those required by their terrestrial counterparts.

As indicated earlier, macrophytes, depending upon their life form among other factors, can absorb nutrients from either the sediment or the water. Algae lack specialized structures for nutrient absorption and therefore derive their nutrition from the water. However, attached algal species are capable of supplementing their nutrition through nutrient absorption from the substratum to which they are attached. This mode of nutrition is particularly common among epipellic and epiphytic forms that obtain some of their nutrients from sediments and plant hosts, respectively. Summarily, the possible modes of nutrition for aquatic primary producers are given below.

| Primary Producer | Nutrition |
|------------------|----------------------|
| Macrophyte | |
| emergent | sediment |
| floating-leaved | sediment and water |
| submergent | sediment and water |
| free-floating | water |
| Algae | |
| phytoplankton | water |
| attached algae | water and substratum |

Both the biomass and the productivity of primary producers are controlled to a large extent by nutrient supply. Accordingly, investigators for many years have sought to predict the growth of primary producers from measurements of essential nutrients. In this regard, nutrient supply has been assessed from measurements of nutrient concentrations in both the environment and in the tissues of primary producer organisms. Additionally, nutrient bioassays have been developed to enable evaluations of nutrient supply from fertilization experiments. All three of these techniques are laden with assumptions and weaknesses; these are discussed briefly below.

In evaluations of water chemistry, the investigator frequently assumes that there is a relationship between concentrations of nutrients in water and the growth of primary producers. Although one can generally expect greater growth in a system where nutrients are abundant, there actually is no simple relationship between the concentration of nutrients present in the system and their availability to the primary producers. The availability of a particular nutrient depends for the most part upon its chemical form, but also upon the ease with which it can be removed from the medium. For example, most of the phosphorus in a sediment is unavailable to rooted plants because it is present in an insoluble form as a component of the sediment matrix. The question of availability has been circumvented to some extent by the development of chemical techniques for determining soluble and extractable ions, which are often assumed to be available to primary producers. Superimposed on the ambiguity associated with determinations of "available nutrients" in the environment, is the question of nutrient supply rate. Theoretically, the supply of a particular nutrient could be adequate at concentrations even near zero, as long as the rate of its replenishment remains equal to the rate of its removal from the environment by

the plant. Certainly, different nutrients are recycled within the environment at different rates. Similarly, primary producers are capable of selective removal of nutrients at different rates. Generally, measurable quantities of nutrients in the environment represent an excess. For these and other reasons, correlations between environmental nutrients and the growth of primary producers are extremely tenuous and need to be interpreted with great caution.

The determination of nutrient concentrations in the biomass (i.e., tissue concentrations) of primary producers provides a more direct means of evaluating nutrient supply. Living tissues have fairly fixed requirements for specific nutrients. At concentrations below the optimal tissue concentration, termed the critical concentration of a particular nutrient, growth of the organism is limited by that nutrient. Conversely, at concentrations above the critical tissue concentration of a particular nutrient, growth of the organism is not limited by that nutrient. Before this technique can be used to evaluate nutrient supply, the critical concentrations of the relevant nutrients must be known. Unfortunately, the process of obtaining this information requires extensive experimentation under carefully controlled conditions. From the literature available on this subject, it appears that critical concentrations vary considerably among different nutrients, plant species, plant parts, and seasons of the year. Because of the great difficulties associated with the separation of algal tissue from other associated organic materials, tissue nutrient analyses in most cases, are not suitable for assessing algal-nutrient supply. In contrast, these analyses are generally quite suitable for assessing macrophyte-nutrient supply; their utility in this regard should greatly increase as more data on macrophyte critical tissue-nutrient concentrations become available.

Nutrient bioassays are essentially fertilization experiments. This technique is based on the premise that nutrient(s) limiting growth will support growth when their concentration in the environment is increased. On one extreme, "whole system" fertilization studies, both in lakes and wetlands, have provided verifiable and important information on the eutrophication process. On the other extreme, short-term "bottle bioassays" have provided confusing and often misleading information when used to predict eutrophication. Short-term nutrient bioassays provide information on potential growth limiting nutrients only during usually very brief and discrete periods of experimentation. Unfortunately, these data are commonly extrapolated to other portions of the year, and in some cases even to reservoirs still in the planning stage. Potentially very serious problems arise from the possibility that nutrients, other than those implicated at the time of the bioassay, may actually exert greater control over the annual growth of the primary producers. Another serious problem related to bottle bioassays is the restriction on the extent of mineral recycling. In nature, recycling occurs among components of the entire ecosystem and not just within the confines of a small

volume of water. Other problems with bottle bioassays can be associated with the use of "test" organisms that often are foreign to the environment being examined. It is extremely important to realize that different primary producers have different nutrient requirements, and thus may actually demonstrate a growth response to different nutrients under the same conditions.

There is a dynamic equilibrium between inorganic and organic components of water. This equilibrium is partially mediated by the activities of primary producers. During the production of biomass, primary producers assimilate inorganic compounds and produce organic matter. When this biomass (organic matter) undergoes degradation, it is reduced to its inorganic components. Thus, the inorganic-organic cycle completes itself. Following periods of active biomass production, nutrients become depleted from the environment and incorporated in the biomass. Contrastingly, following a period of active biomass degradation, the converse occurs. These two processes, of course, occur simultaneously. In addition to changes in concentrations of nutrients and biomass, there are also quite significant changes in the concentrations of dissolved gases associated with these processes. Oxygen supersaturation and carbon dioxide depletion result from high rates of production; rapid degradation results in increased carbon dioxide and diminished oxygen concentrations. The processes of production and degradation respectively promote basic and acidic conditions particularly in poorly buffered and unmixed environments.

BIOLOGICAL INTERACTIONS WITH AND AMONG AQUATIC PRIMARY PRODUCERS

One of the key interactions between algae and macrophytes is their competition for light. For example, submergent macrophytes are disadvantaged because of light limitation in systems where phytoplankton blooms are dense. During blooms, light penetration into the water is severely curtailed; consequently, among the primary producers, only those organisms at or near the water surface can survive. Owing to the marked buoyancy of many species of blue-green algae, these algae are capable of shading other algal species and macrophytes existing at greater depths. In dense macrophyte beds, self-shading by component plants has been demonstrated to affect not only the productivity of the macrophytes, but also the species composition of the macrophyte community.

Competition for nutrients is considered to be an important factor affecting the seasonal distribution and species composition of algae, particularly the phytoplankton. Among the attached algae, competition is also important. For example, it seems reasonable to assume that epiphytic algae compete to some extent with their host for certain nutrients. Similarly, macrophytes may at times compete with the

phytoplankton for nutrients in the water. The extent of this competition would of course, depend on the degree to which the macrophytes derive their nutrients from the water.

Primary producers are affected by various growth-enhancing and growth-inhibiting substances. In general, these substances are dissolved organic compounds resulting from plant excretory and secretory processes. Some types of algae are unable to produce sufficient quantities of certain vitamins for their own use. Algae, primarily those with "animal-like" characteristics, are often dependent upon other primary producers for these materials. Other substances produced by both macrophytes and algae have specific antibiotic effects and thereby can exert some control over community composition. For example, there is limited evidence indicating that certain macrophytes are capable of producing antibiotics having a rather broad range of effectiveness against phytoplankton. Other macrophytes have been demonstrated to produce substances that inhibit the growth of zooplankton, and in some cases, the larvae of mosquitoes as well.

The most obvious interaction between primary producers and animals is grazing. However, the quantity of primary producer tissue consumed in the living state by grazing animals (herbivores) is small, and usually less than 10% is produced. Due to the selective feeding habits of many grazers, the species composition of phytoplankton can be affected. Also, by reducing the overall biomass of phytoplankton and thereby the effects of competition under crowded conditions, grazers may often enhance phytoplankton productivity. The foliage and certain reproductive structures of macrophytes are used as a source of food by a variety of waterfowl and mammals. Although very little macrophyte tissue is directly consumed by grazers, many grazers feed on the epiphytically associated bacteria and algae. The tissue of primary producers is ultimately consumed in the dead state by detritus-feeding organisms. Organisms feeding on dead particulate material in combination with microbial agents of mineralization are ultimately responsible for converting organic matter to its inorganic components.

PRIMARY PRODUCER BIOMASS AND DENSITY

The term biomass refers to the mass of living tissue of an organism or a composite of organisms at a given time. Biomass is commonly presented in units of mass per unit surface area or volume. For example, the biomass of a phytoplankton community can be presented as g/m^2 of water when data on a volume basis are integrated with respect to depth or as mg/l when concentration per unit volume is known. In the strictest sense, and as implicit in the term itself, biomass estimates should be corrected for dead mass. Obviously, this

is often difficult, and in the case of phytoplankton for example, biomass estimates are frequently too large because of the inclusion of suspended dead particulate mass. With macrophyte communities, dead mass is relatively more easily separated from living mass; thus biomass and litter mass categories are routinely distinguished.

Biomass can be determined directly as fresh, oven dry, or organic weight. Fresh weight is not recommended, because it is too variable; but, it is often obtained when large quantities of plant material need to be processed. Oven dry weight is obtained by drying materials in an oven to a constant weight at a temperature of 100°-105° Celsius. Biomass data are most commonly presented in this form. Where the mass of primary producers is affected by inorganic contamination, the determination of weight loss after ignition at 550° Celsius allows the estimation of organic weight in a manner that is unaffected by contamination. In general, the ash content (i.e., mineral portion) of organic matter is less than 10%. When this percentage after combustion is found to greatly exceed 10%, contamination by nonvolatile substances should be suspected. Such a situation commonly arises in hard water lakes where macrophytes become encrusted with carbonate deposits. Also, the belowground biomass of rooted macrophytes is notoriously difficult to wash free of adherent sediment particles. These problems can be circumvented by routinely combusting the plant materials, and expressing the biomass data as organic weight (i.e., ash-free dry weight).

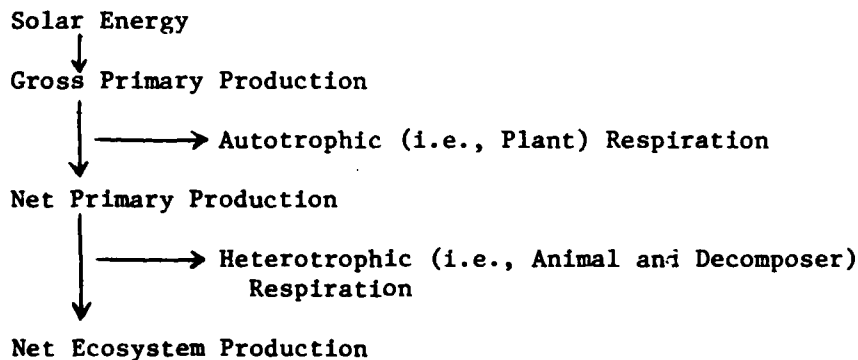
Estimates of algal biomass from chlorophyll or cell volume data are indirect, and should be considered only as approximations. However, because of the previously noted difficulties associated with the separation of algae from detritus, these approximations of algal biomass are commonly used. A few precautions need to be mentioned. Calculating mass from cellular volume assumes constant intracellular density, but in reality intracellular density is quite variable. Also, calculations of cellular volume assume the conformity of algal cells to simple geometric shapes. These shapes need to be carefully matched to different algal species. In regard to estimates of algal biomass from chlorophyll a data, it is important to determine whether or not data have been corrected for chlorophyll degradation products that interfere with actual chlorophyll a measurements. Uncorrected chlorophyll a data are inappropriate for calculating algal biomass.

The density of primary producers is a measure of biomass concentration. More common measures of density are based on number or volume. In turn, these parameters are presented on the basis of either surface area or volume. Numerical density determinations are the most common. With phytoplankton, numbers of algal cells are most frequently expressed as a number per unit volume of water. In the case of attached algae and macrophytes, numerical density estimates are usually given on a surface area of substrate basis. The less commonly determined volumetric density of different algal species can be useful in determining the

relative importance of each individual species within a phytoplankton community.

PRIMARY PRODUCTIVITY

Primary productivity is the rate of production of primary producer biomass. Productivity is expressed as a rate in units of mass (usually dry weight or carbon) per unit area or volume, both per unit time. Sometimes units of energy are substituted for mass, because in the process of primary production, quantities of energy are fixed along with the production of mass. Primary productivity is not measured as a single process. Indeed, the processes of gross primary production, net primary production, and net ecosystem (or net community) production are functionally distinguishable from one another. These processes and their interactions with others are depicted below.



Net Primary Production = Gross Primary Production - Autotrophic Respiration

Net Ecosystem Production = Gross Primary Production - Ecosystem Respiration

Where: Ecosystem Respiration = Autotrophic Respiration + Heterotrophic Respiration

The diagram presented above simplistically demonstrates the relationships of energy and matter in an ecosystem. Gross primary production represents the capacity of primary producers to utilize solar energy in the fixation of inorganic carbon. The organic matter formed by this process provides biological structure and a potential energy base for virtually all other biological processes. Net primary production represents energy and matter produced in excess of the

requirements of the primary producers themselves. These requirements are measured as plant respiration. Net ecosystem production represents energy and matter produced in excess of the requirements of the ecosystems' animals and decomposers. Net ecosystem production can be measured as newly accrued biomass in growing systems, as accumulated detritus in stagnant systems, or as export to adjacent ecosystems.

There are various methods of measuring primary productivity. Among these, the biomass, oxygen, and carbon-14 techniques are the most common. Each of these has inherent strengths and weaknesses, and measures different aspects of the process of primary production.

With the biomass technique, net primary production is determined as the difference between minimum and maximum estimates of biomass. This method is appropriate only when interim losses of biomass due to grazing, mortality, or other sources are negligible or can otherwise be accounted for. For this reason, the technique cannot be applied to algae. It is best applied to macrophyte populations that demonstrate distinct annual periods of growth. Determinations of the productivity of perennial macrophyte populations are often problematical, because biomass often remains unchanged, and also since annual patterns of growth are confused by variable rates of mortality and internal translocations of mass above and below ground.

The oxygen technique involves measurements of community respiration in the dark (O_2 loss) and net community production (O_2 gain) in the light. The sum of these two quantities provides an estimate of gross primary production. This technique does not measure net primary production, because plant's own respiration is inseparable from other sources of respiration such as microbial decomposition. This technique is applicable to the phytoplankton for determinations of gross production, net community production, and pelagic (i.e., open water) respiration. In general, the oxygen technique is not applicable to macrophytes, because of their ability to internally store and recycle gases including oxygen. Thus macrophytes are capable of obscuring changes in oxygen concentration associated with internal processes of production and respiration.

Among the various methods of determining phytoplankton productivity, the carbon-14 technique is becoming the most commonly used. It is advantageous because of its simplicity and directness. Rather than measuring changes in concentrations of substrates or products of photosynthesis as is done with other techniques, the C-14 technique directly measures carbon fixation. This measure is achieved by monitoring the photosynthetic incorporation of radioactively labeled inorganic carbon (i.e., carbon-14). Because some of the fixed carbon is lost to the plant's own respiration, the C-14 technique is thought to measure net primary production. However, because this loss appears to be minimal under certain conditions, this technique can also be considered to measure gross primary production. For practical purposes, investigators usually assume that the C-14 technique measures something

between gross and net primary production. Whatever is the case, this technique affords excellent replicability and is extremely sensitive. Carbon-14 has been used with some success to measure the productivity of attached algae as well as phytoplankton. Although the technique can also be adapted to measurements of macrophyte productivity, such applications are methodologically cumbersome, and the results are difficult to interpret.

CONCLUSIONS

In this article I have briefly reviewed and summarized the more important aspects of the role of primary producer organisms in aquatic ecosystems. Much of the information contained herein is of quite recent origin, and consequently subject to change as new data become available.

Primary producer organisms are an integral part of the aquatic environment and exert a significant influence on the quality of the same. If Federal agencies and the public in general are seriously concerned about the degradation of water quality in this country, then the research efforts of qualified laboratories and individuals must continue to be supported. The importance of sustained research related to the biology and role of primary producers in aquatic ecosystems cannot be overemphasized.

ACKNOWLEDGEMENTS

Joe H. Carroll, Douglas Gunnison, and R. Michael Smart of the Ecosystem Research and Simulation Division of the Environmental Effects Laboratory are gratefully acknowledged for critically reviewing this article. Gordon L. Godshalk of the Department of Botany, University of Florida, assisted in a similar capacity, and his help is most appreciated. Thanks also to our division's secretarial staff for their assistance in the preparation of this manuscript.

SELECTED BIBLIOGRAPHY

- Cooper, J. P. (ed.). 1975. Photosynthesis and productivity in different environments. Cambridge University Press, New York.
- Evans, L. T. 1963. Environmental control of plant growth. Academic Press, New York.

- Gauch, H. G. 1973. Inorganic plant nutrition. Dowden, Hutchinson, and Ross, Inc., New York.
- Gessner, F. 1955. Hydrobotanik. Vol. 1. Energiehaushalt. VEB Deutscher Verlag der Wissenschaften, Berlin.
- Gessner, F. 1959. Hydrobotanik. Vol. 2. Stoffhaushalt. VEB Deutscher Verlag der Wissenschaften, Berlin.
- Goldman, C. R. (ed.). 1965. Primary productivity in aquatic environments. Proceedings of an IBP-PF Symposium. Memorie dell'Istituto Italiano di Idrobiologia 18 Suppl. (Reprinted 1966; same title, University of California Press, Berkeley.)
- Hutchinson, G. E. 1957. A treatise on limnology. Vol. I. Geography, physics, and chemistry. John Wiley and Sons, Inc., New York.
- Hutchinson, G. E. 1967. A treatise on limnology. Vol. II. Introduction to lake biology and the limnoplankton. John Wiley and Sons, Inc., New York.
- Krebs, C. J. 1972. Ecology, the experimental analysis of distribution and abundance. Harper and Row, Inc., New York.
- Kusnetsov, S. I. 1959. Die Rolle der Macroorganismen in Stoffkreislauf der Seen. VEB Deutscher Verlag der Wissenschaften, Berlin.
- Lieth, H. and R. H. Whittaker (ed.). 1975. Primary productivity of the biosphere. Springer-Verlag, New York.
- Lewin, R. A. (ed.). 1962. Physiology and Biochemistry of algae. Academic Press, New York.
- Odum, E. P. 1971. Fundamentals of ecology (3rd Edition). W. B. Saunders Company, Philadelphia.
- Round, F. E. 1965. The biology of the algae. E. Arnold Publishers, Ltd., London.
- Ruttner, Franz. 1972. Fundamentals of limnology (3rd Edition). University of Toronto Press, Toronto.
- Sculthorpe, C. D. 1967. The biology of aquatic vascular plants. St. Martin's Press, New York.
- Stein, J. R. (ed.). 1973. Phycological methods. Cambridge University Press, London.
- Stewart, W. D. P. (ed.). 1974. Algal physiology and biochemistry. University of California Press, Berkeley.
- Vollenweider, R. A. (ed.). 1969. A manual on methods for measuring primary production in aquatic environments. Int. Biol. Program Handbook No. 12. Blackwell Scientific Publications, Oxford.
- Wetzel, R. G. 1975. Limnology. W. B. Saunders Company, Philadelphia.
- Woodwell, G. M. and E. V. Pecan. 1973. Carbon and the biosphere. Technical Information Center, Office of Information Services, AEC, Washington.

WATER QUALITY EFFECTS OF MULTI-PURPOSE RESERVOIRS AND NAVIGATION DAMS ON THE MONONGAHELA RIVER

by

Michael Koryak¹

I. INTRODUCTION

This is a case study of a basin-wide system of Corps of Engineers water resource projects. The discussion emphasizes functional aspects of the data and how it is being utilized to establish design and operational criteria.

II. DESCRIPTION OF THE AREA

The Monongahela River is formed by the confluence of the West Fork and Tygart Rivers at Fairmont West Virginia (Plate 1). From Fairmont, the Monongahela River flows in a northerly direction for 128.7 miles before joining the Allegheny River at Pittsburgh, Pennsylvania to form the Ohio River.

The river drains 7,386 square miles of northern West Virginia, western Maryland, and southwestern Pennsylvania.

Slackwater navigation is maintained year-round on the entire 128.7 mile length of the Monongahela River by a series of Corps of Engineers low-head navigation dams. There are nine navigation Locks and Dams (L/D) on the mainstem of the Monongahela River while the lower 11.2 miles are in the pool of Emsworth L/D, located at Ohio River mile 6.2.

There are two large tributary storage reservoirs in the basin, Tygart River Lake and Youghiogheny River Lake. Also, one additional reservoir, Stonewall Jackson Lake, is authorized on the West Fork River.

The navigation system links the coal fields of northern West Virginia and southwestern Pennsylvania with the basic metal industries of the Pittsburgh region. Consequently, acid mine drainage from bituminous coal mining activities is the principal water quality problem in the upper basin, and industrial, thermal, and domestic pollution degrade the lower 30 miles.

The Monongahela River is very low yielding at base flow and most water quality problems are low-flow related.

¹ Aquatic Biologist, Hydrology and Hydraulics Branch, U.S. Army Engineer District Pittsburgh.

III.A. TYGART RIVER LAKE

Tygart Dam is located in northern West Virginia, and controls 1,184 square miles of the Tygart River Basin. At summer pool, the lake is 10 miles long with maximum and mean depths of 134 feet and 63 feet respectively. By May of each year there is usually 110,000 acre feet of water stored, and all but 9,700 acre feet of this total is available for augmentation.

There is no authorized water quality storage in Tygart River Lake. The official project purposes are flood-control and low-flow augmentation to maintain uninterrupted navigation on the Monongahela and upper Ohio Rivers. We can and do however, utilize excess storage from Tygart River Lake for water quality purposes.

III.B. AUGMENTATION FROM TYGART DAM TO MITIGATE ACID MINE DRAINAGE

Tygart River Lake provides an assured minimum flow of 340 cfs at Opekiska L/D near the head of the Monongahela River mainstem. We also attempt to maintain a 2 to 1 flow ratio between the discharge from Tygart Dam and the flow of the mine drainage degraded West Fork River.

Tygart Dam has been operational since 1939 and it's somewhat difficult to evaluate the effectiveness of an operation that has been going on continuously for almost forty years. Plate 2 shows the reduction in summer acidity values at McKeesport, Pennsylvania that occurred when operations to augment Monongahela River flows were initiated in 1939. These values were measured at a point 16 miles above the mouth of the Monongahela River and 135 miles downstream of Tygart Dam.

Another factor to be considered in the evaluation of the data presented in Plate 2 is the fact that during the depression years, prior to the construction of Tygart Dam, coal production in the basin was very low. With the beginning of World War II, shortly after completion of the project, coal production in the Monongahela River Basin increased substantially. Therefore the reduction in summer acidity values shown in Plate 2 occurred during a period when increased acidity would have been expected.

III.C. INFLOW QUALITY

Bituminous coal mining is also an important economic activity upstream of Tygart Dam, and periodic intrusions of acid mine drainage enter Tygart River Lake during the summer and autumn months.

A pattern of flow related, depressed inflow pH values during the summer of 1973 can be seen in Plate 3. The outflow pH is also plotted in Plate 3 for the same period. In spite of the extreme, periodic degradation of the inflow water a high quality discharge is consistently maintained at Tygart Dam.

Biological samples were collected that verify this observation. Very low macroinvertebrate productivity at the inflow station indicated stress from acid mine drainage. At the outflow station a 5,000% increase in the average macroinvertebrate dry weight per square foot was observed (from 1.2 mg/ft² to 62.2 mg/ft²). The inflow samples were numerically dominated by acid tolerant Chironomidae (53.8%). At the outflow, the percent composition of Chironomidae dropped to 32.3%, while the more sensitive Ephemeroptera increased from 0 to 66.1%.

Also consistent with these observations there is an excellent outflow sports fishery at Tygart, and no fishery at all at the inflow.

III.D. INTERNAL HYDRODYNAMICS AND WITHDRAWAL ELEVATION

The depth of penetration and resulting mixing patterns of the acid inflows into Tygart River Lake are influenced by the design and operation of the dam. The internal hydrodynamics of the reservoir in turn influence the chemistry and biology of both the impoundment and the outflow.

Tygart River Lake has bottom withdrawal, and a relatively short hydraulic retention time. The volume of water normally discharged between June and August is equivalent to 1.65 times the storage at summer pool. In spite of only moderate vertical thermal gradients, inflows can penetrate the impoundment as well-defined temperature-density currents. Interflow occurs during most of the summer months when acid loading problems are most severe.

A seasonal cycle of temperature-density current variation can be seen in Plate 4. Conductivity values are used in this plate to follow the path of the density flows that originate from the Tygart River inflow. Beginning in May, the impoundment is still cool, nearly isothermic, and conductivity values are low and relatively uniform throughout the reservoir. During this time of year, the more easily warmed inflow can be expected to overflow the cooler denser reservoir as seen on the 26 June 1973 survey.

With summer stratification, interflow conditions gradually develop. Interflow is first clearly detectable on the basis of conductivity values in early July in the upper reservoir. This interflow penetrated Tygart River Lake in a horizontal layer concentrated at approximately 30-40 feet below the surface.

There was constant mixing and entrainment, especially in the shallow upper portion of the impoundment. But for the most part the current retained its identity and could still be recognized when it reached the dam, 10 miles downstream and roughly 30 days after first entering the impoundment.

The conductance models were generally useful in describing mixing patterns of less conservative acid mine drainage related parameters. For example, Plate 5 demonstrates a close correlation between pH and conductance distribution patterns for a 15 August 74 survey.

Because of bottom withdrawal, sharp slugs of acid that move through Tygart River Lake as either overflow or interflow are will mixed and diluted before they are discharged from the Dam. This accounts for the relative consistency in the quality of the outflow. In the autumn, the inflows are likely to cool before the reservoir, and underflow conditions can develop, as was noted in late September 1973 (Plate 4). Underflows short circuit the reservoir. With bottom withdrawal the retention time of underflows is short, and mixing and dilution is not as thorough as with overflow and interflow.

Selective withdrawal options are not available at Tygart River Lake, but if they were, a higher elevation discharge would probably only be desirable during underflow periods. During these periods the higher level withdrawal would protect discharge water quality.

Considering the nature of the problem at Tygart, the low level outlets of the existing structure generally provide an acceptable means for the partial mitigation of acid slugs in the outflow without sacrificing reservoir water quality and the reservoir fishery. Withdrawal from a higher elevation during the summer months would not only create downstream pH depressions, but would result in concentrated acid interflows penetrating the reservoir at a higher elevation. This would have a more negative impact on the biologically sensitive surface strata.

IV.A. YOUGHIOGHENY RIVER LAKE

Youghiogheny Dam is located 74 miles above the mouth of the Youghiogheny River. The Youghiogheny River makes its confluence with the Monongahela River 16 miles upstream of Pittsburgh, Pennsylvania. Youghiogheny Dam controls a 434 square mile drainage area. Total storage at summer pool is 154,000 acre-feet and the project has been totally operational since 1948.

Unlike Tygart Reservoir, there is authorized water quality storage in Youghiogheny River Lake, and there are no significant sources of acid mine drainage upstream of this project. Similar to Tygart, however, the dilution and neutralization of downstream acid mine drainage is an important aspect of the operation of Youghiogheny River Lake.

IV.B. AUGMENTATION FROM YOUGHIOGHENY RIVER LAKE

Plate 6 shows the pH of the Monongahela River above and below the confluence of the Youghiogheny River in 1965. Except for a very brief period, downstream pH was higher than at the upstream station, indicating the year round positive influence of the Youghiogheny River. During the summer months, the pH of the station above the confluence of the Youghiogheny River dropped dramatically while the pH of the station downstream of the Youghiogheny River actually climbed during this same period. No major sources of alkalinity other than augmented Youghiogheny River flows could be identified in the reach between the two stations.

As illustrated by the data presented in Plate 6, the Youghiogheny River Lake water quality augmentation program has been highly successful. One limitation to effective year round water quality control however, and a potential constraint in any water quality augmentation operation, is also evident in this plate.

In April 1965 the pH of both Monongahela River stations was mildly depressed. Similar spring pH depressions also occur regularly in both the Tygart and Youghiogheny Rivers. The timing of these pH depressions corresponds with the spring filling cycles of the tributary storage reservoirs. Excess run-of must be stored during the spring for low-flow augmentation during the dry summer and autumn months when severely acidic conditions are most likely to develop in the downstream target areas. During the filling cycles, and when flood control operations are made, however, the percentage flow contributions of uncontrolled polluted tributaries can be quite high.

Acidity values in the acid mine drainage effected streams of the upper Ohio River Basin are generally inversely related to flow. Therefore, the filling cycle problems are always relatively mild and of short duration compared to probable summer conditions without augmentation.

Thermal pollution and anoxia also exhibit a strong inverse relationship to flow. Plates 7 and 8 show the mitigating effect of the cool, well aerated Youghiogheny River on thermal pollution and anoxia respectively in the lower Monongahela River during a 1975 summer low-flow survey. The flow in the Monongahela River upstream of the Youghiogheny confluence at this time was 700 cfs. The flow in the Youghiogheny was 1000 cfs, 75% of which was augmented flow from Youghiogheny River Lake.

V. MONONGAHELA RIVER NAVIGATION STRUCTURES

Navigation structures on the Monongahela River directly influence a number of physical, chemical, and biological parameters. Dissolved oxygen is the most significantly affected water quality parameter. The degree and nature of the effect is dependent on the design and operation of the individual structures, localized biochemical oxygen demands, and flow and temperature conditions.

At high flows turbulent reaeration can be expected below all of the structures. Problems can occur however at low summer flows, and because of some design features, particularly downstream of Opekiska L/D.

Dissolved oxygen stratification can develop in the larger navigation pools at low summer flows. The dissolved oxygen concentration of Opekiska Pool during the July 1975 survey (Q 650 cfs) presented in Plate 8 decreased vertically from a maximum of 10.2 mg/l at the surface to a minimum of 0.1 mg/l near the bottom.

The invert elevation of the outlets of Opekiska Dam (fixed sill elevation 831 msl), is 26 feet below the surface of Opekiska Pool (normal pool elevation 857 msl). At this depth, during low summer flow conditions, the concentration of dissolved oxygen is low. The outlets of Opekiska Dam are submerged in Hildebrand pool (normal pool elevation 835 msl) and the low-flow reaeration potential of such an arrangement is limited. As illustrated in Plate 8, the Opekiska discharge can depress dissolved oxygen concentration over the entire length of Hildebrand Pool. Similar D.O. depressions have also been observed below other submerged sill, gated dams on the mainstem Ohio River at low flow.

A stratified pattern of dissolved oxygen distribution was also apparent in Hildebrand Pool during the July 1975 low-flow survey. The vertical dissolved oxygen profile varied from a maximum of 6.4 mg/l at the surface to a minimum of 3.9 mg/l at a depth of 29 feet. The elevation of the "hypolimnetic" strata of Hildebrand Pool appeared to be influenced by the invert elevation of the outlets of Hildebrand Dam (816 msl or 19 feet below the normal pool elevation).

The outlets of Hildebrand Dam are elevated above the next receiving pool, Morgantown Pool, and turbulent reaeration from Hildebrand Dam was sufficient to mitigate the dissolved oxygen depression that was originally created by the Opekiska Dam discharge. Reaeration at Hildebrand Dam may also be increased by bucket sills which flip the discharged water.

Similar to what was observed in Hildebrand Pool, Morgantown Pool appeared to trend towards low-flow dissolved oxygen stratification below the invert elevation of the outlets of Morgantown Dam. However, this pattern in Morgantown Pool was weakly developed.

Some low-flow reaeration was noted in L/D 8 Pool from the discharge of Morgantown Dam. Morgantown Dam also has flip bucket sills. Likewise below L/D 8 some reaeration was measureable at low-flow. The sills at L/D 8 are elevated 12 feet above normal L/D 7 Pool.

In general, with the exception of Opekiska L/D, either limited reaeration or no significant change was observed below all mainstem navigation structures on the Monongahela River.

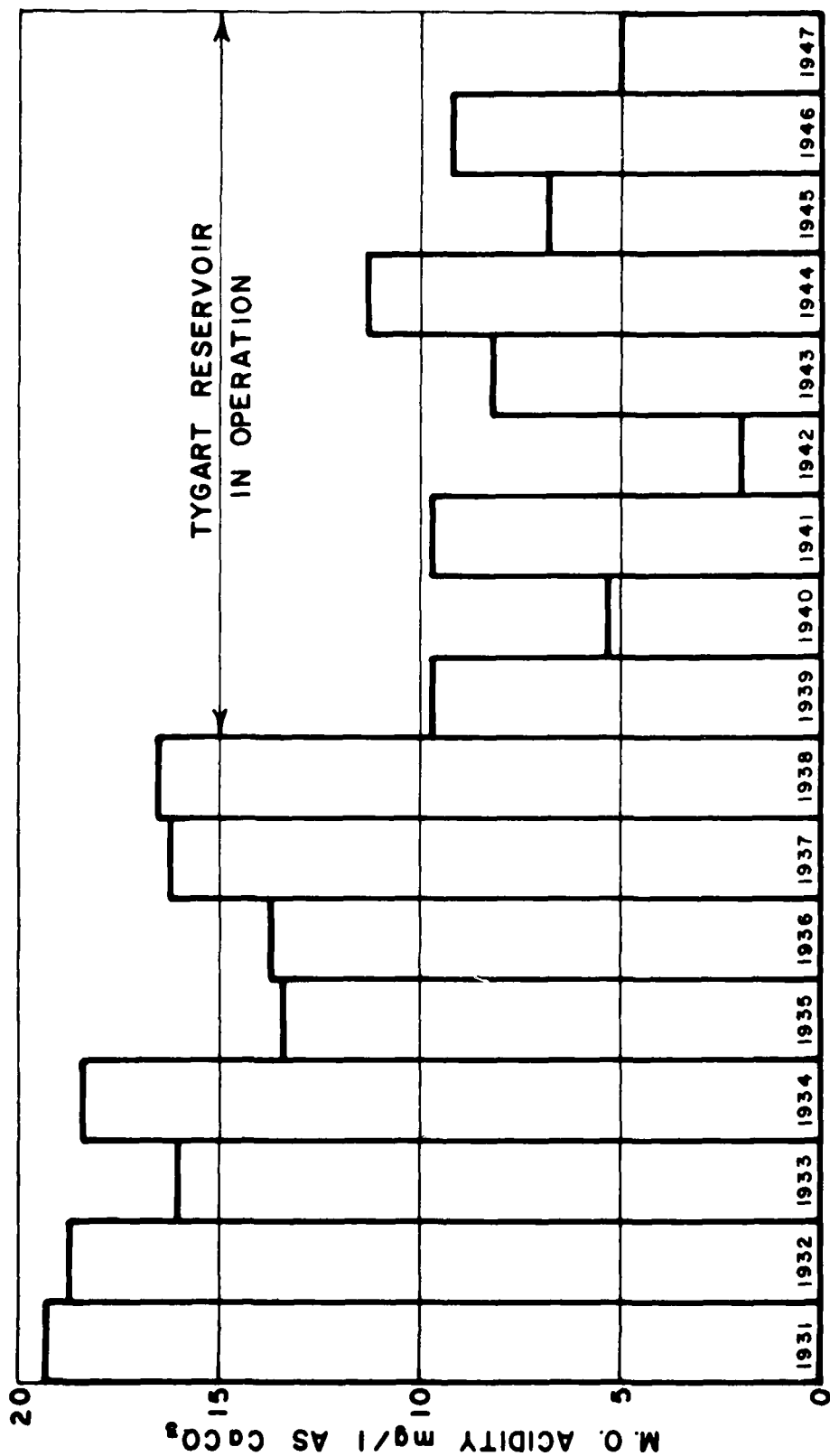
Existing gated navigation projects are being re-examined to determine the feasibility of increasing aeration at low-flow by the adjustment of the current operation of the gates.

Since it has been established that these structures aerate when discharging high volumes of water, it is likely that further increasing the discharge per gate at low flows by using fewer gates with larger openings may increase low flow turbulence and aeration.

We have already had some success with operationally induced turbulence on the mainstem Ohio River. Hydraulic studies at Montgomery L/D for instance, demonstrated that 1 gate opened 2 feet provided $1\frac{1}{2}$ times the amount of D.O. as 2 gates opened 1 foot.

Even where it may be possible to successfully operate these structures for reaeration, the restrictions imposed by stilling basin action, downstream riprap stability, currents, channel scour, available tailwater, etc., may limit operations to create turbulence for the sake of reaeration alone.

The problem can be approached more effectively in the design of new facilities. Currently major portions of a modernization program for the Monongahela River navigation system are in the planning stage. The program will most likely consist of replacement structures for L/D 7 and L/D 3, and structural modifications at L/D 2, L/D 4, and L/D 8. Low-flow dissolved oxygen conditions in the lower river indicate that significant water quality benefits could be realized at low additional costs by designing and incorporating aerating hydraulic features into L/D 2 and L/D 3.



MONONGAHELA RIVER ABOVE MCKEESPORT, PA.
 AVERAGE METHYL ORANGE ACIDITY (JUNE - NOVEMBER)

AD-A101 802

CORPS OF ENGINEERS WASHINGTON DC
PROCEEDINGS OF A SEMINAR ON WATER QUALITY DATA INTERPRETATION, --ETC(U)
1978

F/G 13/2

NL

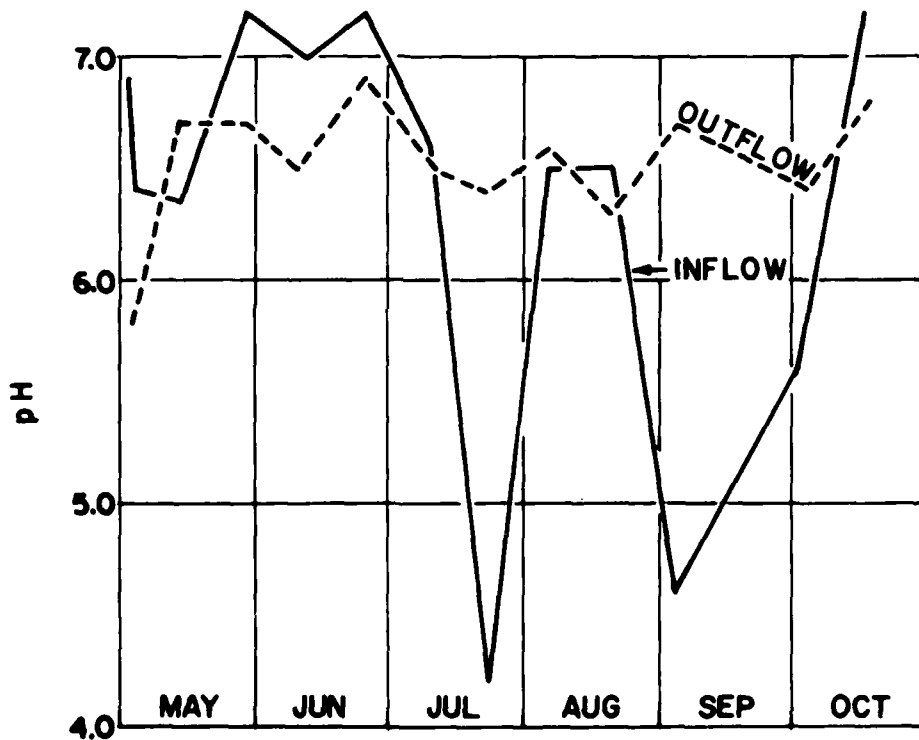
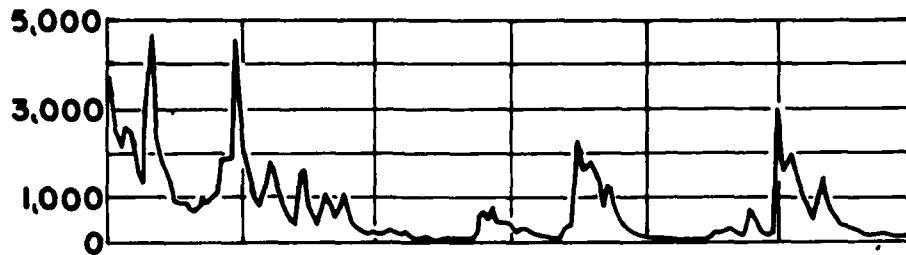
UNCLASSIFIED

3-3
DTIC

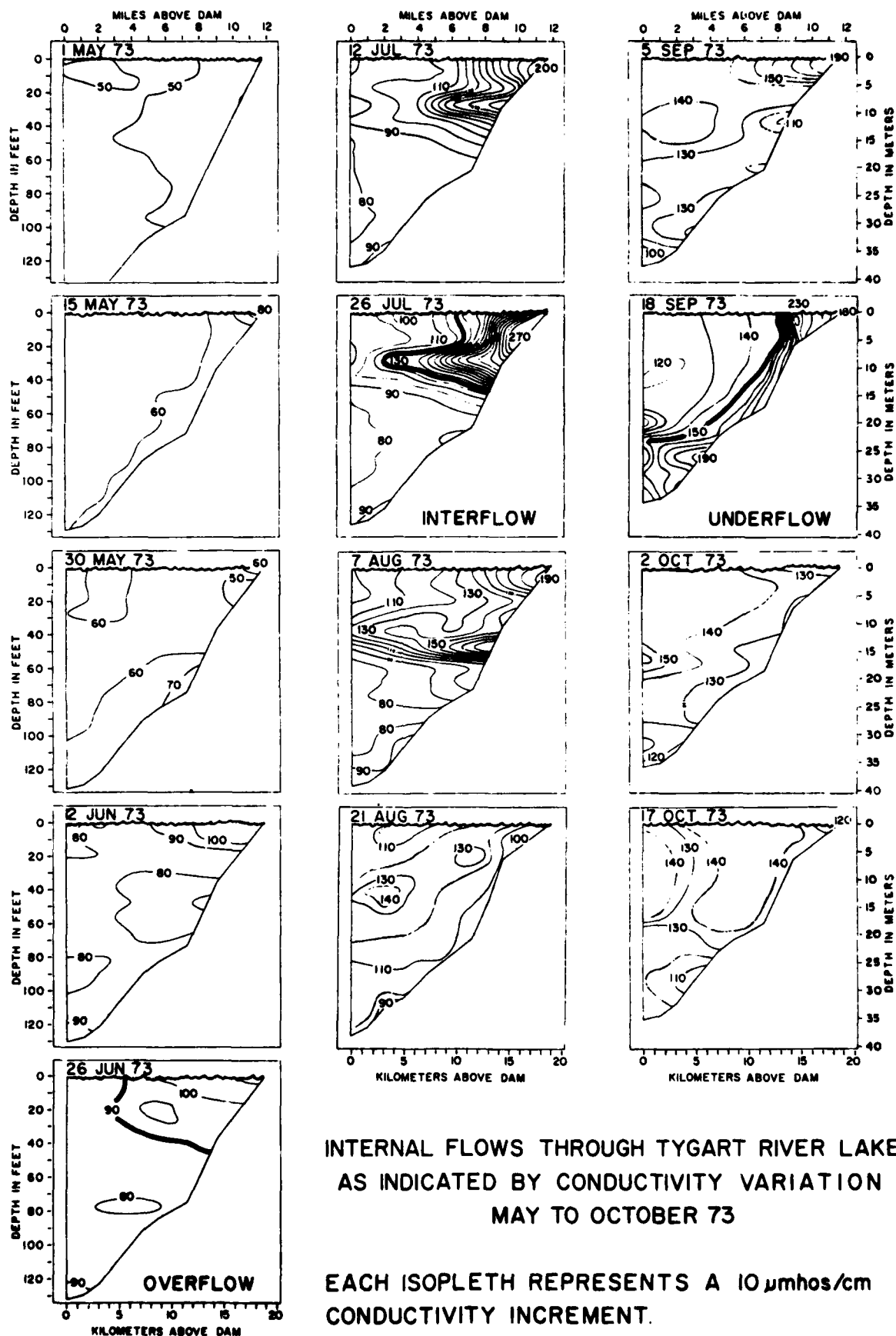


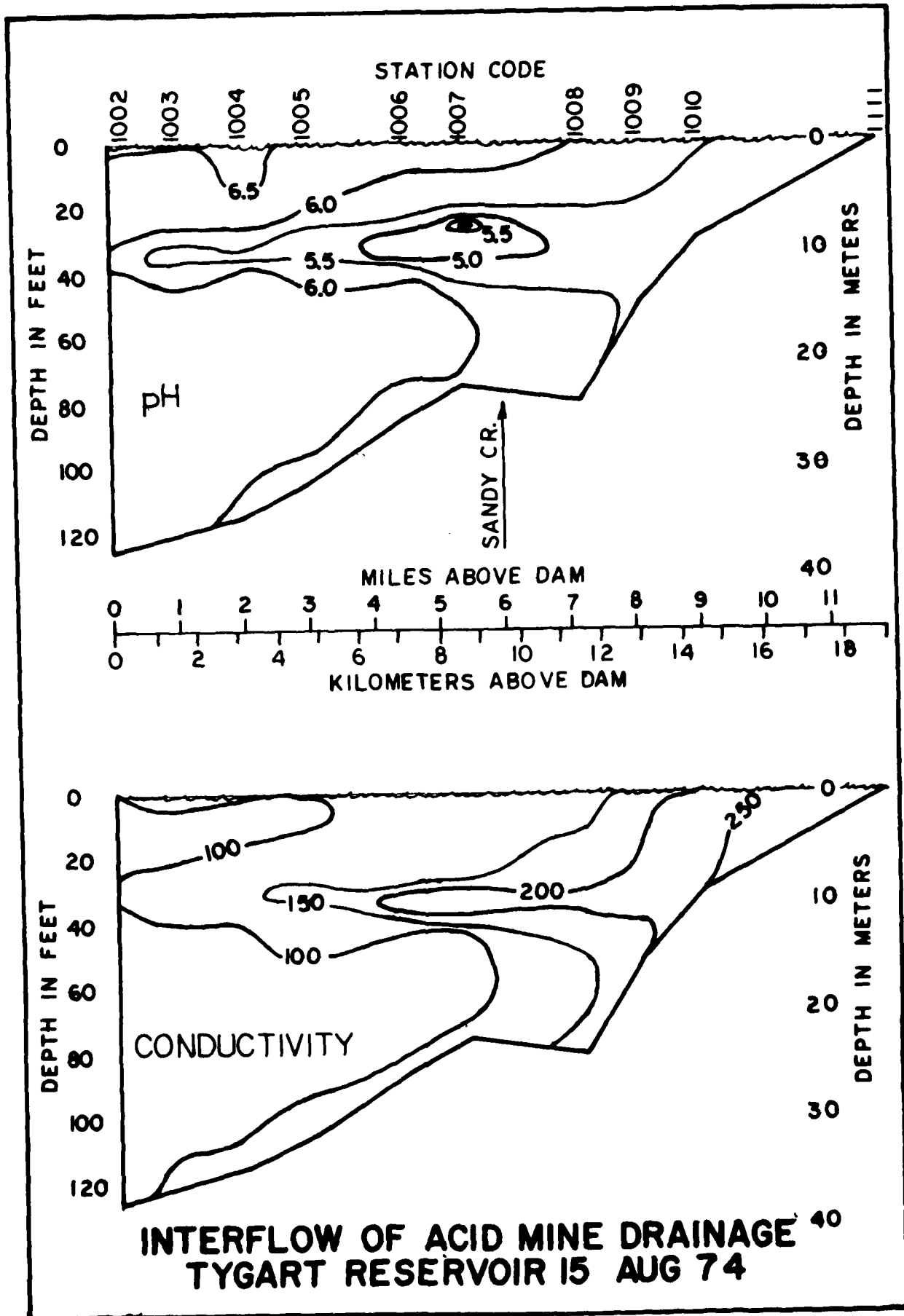
END
DATE
FILMED
8-8-81
DTIC

PHILIPPI GAGE
MEAN DAILY FLOW
IN C.F.S.



TYGART RIVER LAKE 1973
INFLOW AND OUTFLOW pH

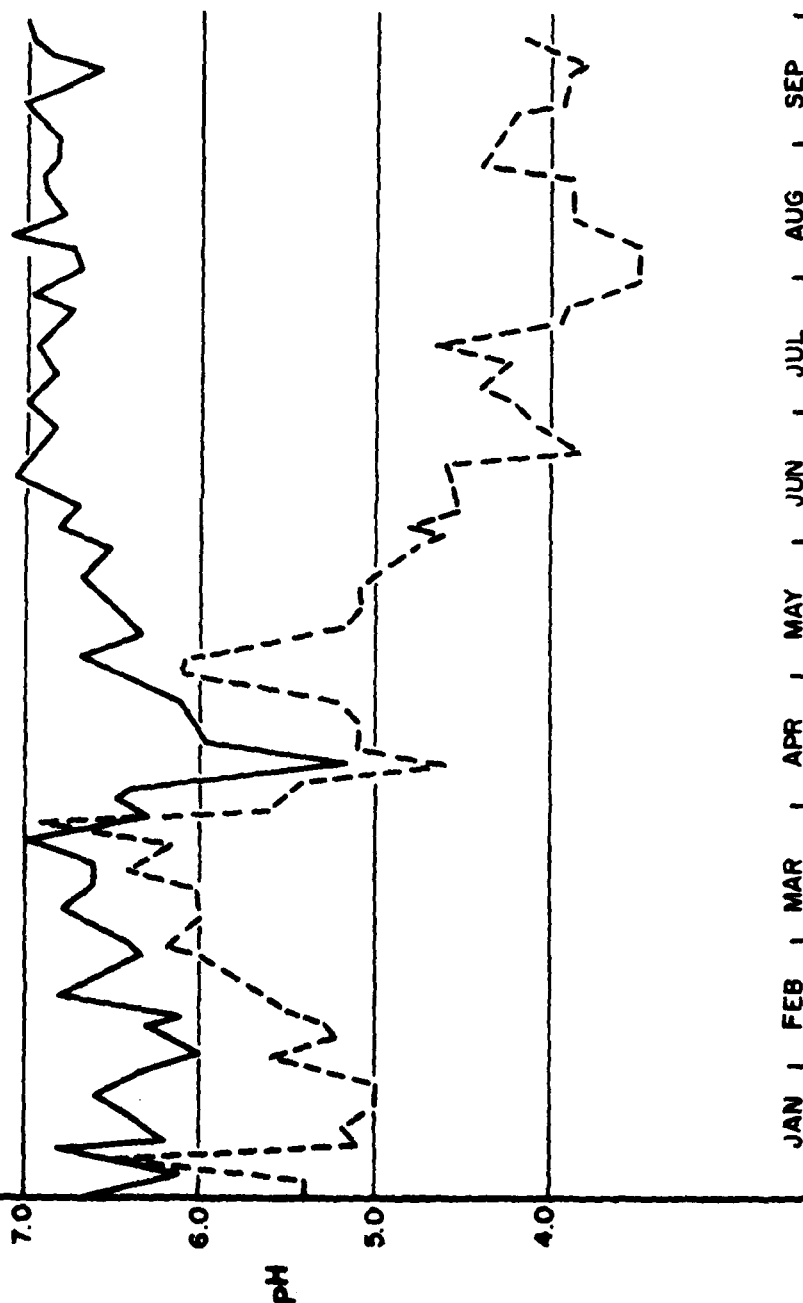




MONONGAHELA RIVER pH 1965

DOWNSTREAM OF THE
MOUTH OF THE
YOUGHIOGHENY RIVER
RIVER MILE 0.6

UPSTREAM OF THE
MOUTH OF THE
YOUGHIOGHENY RIVER
RIVER MILE 43.3



*EACH ISOPLETH REPRESENTS A 1.0°C WT. INCREMENT
 *THE LOCATIONS OF INDUSTRIES WITH AN AVERAGE
 WITHDRAWAL OVER 200 MGD ARE INDICATED.

